UNIVERSITEIT GENT

Faculteit Psychologie en Pedagogische Wetenschappen

Lexical and Semantic Organization in Bilinguals

Wouter Duyck

Promotor: Prof. Dr. André Vandierendonck
Co-Promotor: Prof. Dr. Marc Brysbaert

Proefschrift ingediend tot het behalen van de academische graad
van Doctor in de Psychologische Wetenschappen

2004

UNIVERSITEIT GENT
Faculteit Psychologie en Pedagogische Wetenschappen

Lexical and Semantic Organization in Bilinguals

Wouter Duyck

Promotor: Prof. Dr. André Vandierendonck
Co-Promotor: Prof. Dr. Marc Brysbaert

Proefschrift ingediend tot het behalen van de academische graad
van Doctor in de Psychologische Wetenschappen

2004

CONTENTS

ACKNOWLEDGEMENTS

This doctoral thesis was accomplished while I was working as a Research Assistant of the Fund for Scientific Research – Flanders (Belgium) (F.W.O. – Vlaanderen) at the Department of Experimental Psychology of Ghent University, Ghent, Belgium. I would like to thank these institutions for their support.

First, I would like to thank two great researchers, who have introduced me in the field, and who have been inspiring examples ever since. Prof. Dr. André Vandierendonck was the first to arouse my interest for human cognition. He convinced me that psychological science could advance much through experimental research. I want to thank him for being the promotor of this dissertation during the last two years of the project. It was much appreciated. I also would like to thank Prof. Dr. Marc Brysbaert, the first promotor of this project, who introduced me into psycholinguistics, shortly after my first steps in experimental psychology. His incredible drive and expertise in various domains set an example difficult to match; his e-mails were always motivating. I want to thank both for the freedom they offered me during this project.

Second, I would like to thank my parents for the opportunity to study in ideal circumstances. I have great respect for the way they succeeded in life without such support.

Special thanks go to two four-star friends (and colleagues). Arnaud Szmalec has been a true friend during the last years. I would like to thank him for being him and for teaching me the Polish national anthem. Secondly, I would like to thank Wouter De Baene, for putting up with me four years in the same office, for the entertainment there, for outperforming any spelling checker on every manuscript you send him within the hour (though he missed the double point at the end of this sentence)..

Furthermore, I thank Timothy Desmet, an excellent psycholinguist, for being a great discussion partner (from soccer to Pinker) and for making the 'difficult decisions' in the Wines and Bubbles bar. I would also like to thank all the players of my soccer team (from which I would like to avoid mentioning the name, as it contains the letters 'w', 'a', 'c', 'k' and some vowel). Thanks also to my other team, 'Standard' Bellegem, for the warm reception, for the anthropologically interesting atmosphere, for walking the thin line between FC De Kampioenen and Real Madrid. Thank you to Fitus, for the extra 10 pounds her cookery caused me, for the nice holidays.

The work presented in this thesis also benefited from the many different psycholinguists and other scientists I had the pleasure to work with: Frederik Anseel, Wendy De Moor, Timothy Desmet, Kevin Diependaele, Denis Drieghe, Rob Hartsuiker, Dominiek Sandra and Sofie Schoonbaert and Lieven 'WordGen' Verbeke. For excellent technical assistance, I am grateful to Pascal Mestdagh and Antoine Tavernier. Also, I would like to thank Lies Baten and Christophe Blanckaert.

Finally, I would like to thank Kielie, for tolerating the laptop which was my third arm during these last weeks, for listening to my work-related nagging, for being the most wonderful person I've ever met, for that one special gift.

Wouter Duyck

Ghent, February 10th, 2004

CHAPTER 1
BILINGUAL LANGUAGE ORGANIZATION: AN OVERVIEW[1]

Without bilinguals, the world would be a modern Tower of Babel. Almost everybody has ever experienced the discomfort of having to rely on primitive communication methods (e.g. gestures), being in a country where nobody speaks a language you know. Luckily, about half of the world's population has reasonable knowledge of more than one language and, thus, can be considered bilingual (Grosjean, 1982, pp. vii). This estimate further increases if one regards significant differences between home dialects and official languages as another form of bilingualism. After all, the fact that most dialects are not treated as separate languages is politically motivated rather than scientifically based (Fabbro, 1999). Furthermore, widespread bilingualism is not a privilege of 'developed' Western countries such as notorious bilingual Belgium (Dutch-French), Canada (French-English) or Wales (Welsh-English). In Cameroon for example, there are two official languages, four 'lingua francas' and 285 dialects. More than half of the population speaks three or more of these languages (Bamgbose, 1994). From the above, it will be clear that the term 'bilingual' in the present dissertation (as in the psycholinguistic literature) does not only apply to people who speak two languages equally well because their parents had two different native languages for example. Instead, in the present dissertation, "*Bilingualism is the regular use of two (or more) languages, and bilinguals are those people who need and use two (or more) languages in their everyday lives.*" (Grosjean, 1992, pp. 51). Finally, it is worth noting that it does not hurt to be a bilingual. In contrast with the earlier widespread belief that being bilingual causes mental retardation (for a history, see Hakuta,

---

[1] Parts of this chapter are based on the introductory section of an article by Duyck and Brysbaert (2002).

1986), bilingualism has been associated, among others, with higher metalinguistic abilities (Hakuta & Diaz, 1985) and divergent thinking capacities (Landry, 1974).

In contrast to the omnipresence of bilingualism, psycholinguistic research on the phenomenon is relatively rare and has developed only recently. Most influential models of visual word recognition (e.g. McClelland & Rumelhart, 1981) and speech production (e.g. Levelt, 1989) are still exclusively monolingual models, making the implicit assumption that bilinguals are in fact two monolinguals in the same person. From the present dissertation, it will become clear that this assumption may not be justified, as representations of both languages constantly interact with each other. It is not the case that people can just ‘switch off’ their first language, or even their second. Therefore, I agree with De Bot (1992) that monolingual models of language processing should actually be special cases of more general, bilingual models. From this perspective, research on bilingualism is also important for the literature on monolingual language processing.

This introductory chapter provides the general reader with a broad overview of the major findings in the psycholinguistic literature on bilingualism. This review, like the dissertation itself, is mostly confined to the processing of visually presented words. Second, I will describe three influential models of bilingualism which have been successful in explaining the major empirical findings. Finally, the research objectives are stated and an outline of the experimental studies that are presented in this dissertation is given.

INTRODUCTION: SEPARATE LEVELS OF LANGUAGE REPRESENTATION

Essentially, to become bilingual, one must acquire the capacity to derive meaning from second language word forms (for listening and reading), and the capacity to produce meaning with these new forms (for speaking and writing). In addition, the meaning expressed by words from the second language (L2) is likely to be closely related to the meaning that otherwise would be conveyed with words from the first language (L1), even though the word forms of both languages may be very different. The fact that two different word forms are mapped on the same semantic concept in bilinguals is one of the reasons why researchers have started to think of visual word recognition as a process involving at least two different kinds of representations. The first representation has to do with the word forms and is generally called the lexical level (because the 'dictionary' of known words is referred to as the mental lexicon). The second representation is related to the meaning of the words and is called the semantic level[2].

As noted by Kroll (1993), many contradictory findings in early research about the organization of the bilingual language system originated from the fact that researchers of bilingualism did not make a clear distinction between lexical and semantic word representations. Studies that emphasized word meanings mostly produced evidence for a single language system shared by both languages, whereas studies that primarily addressed lexical processes seemed to provide support for two distinct, language-specific systems. In the section below, I will first review the evidence in favor of a single semantic system accessed by L1 and L2 words, and then address the issue of how the lexical (orthographic) level of a bilingual should be thought of. Finally, I will discuss a third representational level, that of phonological

[2] Note that the distinction between lexical and semantic information does not imply serial processing, in which a word form must first be identified *before* meaning can be attached. Most current models of word recognition see lexical activation as a competition process that takes some time and during which activation continuously dissipates from the lexicon to the semantic system (and, according to some models, returns due to top-down connections).

representations[3], which has received almost no attention in the bilingual literature until now.

THE ORGANIZATION OF SEMANTIC REPRESENTATIONS IN BILINGUALS

There are several sources of evidence that L1 and L2 words access a common conceptual system. First, studies of interference effects, such as the negative priming effect, have repeatedly shown that processing in one language interferes with processing in the other language (see Francis, 1999, for a review). For instance, Fox (1996) presented English-French bilinguals with two displays per trial. On the first display, an Arabic digit was shown together with the same word printed above and beneath the numeral (e.g. the digit 5 between the words *pepper* and *pepper*). Participants were asked to indicate whether the digit represented an odd or an even number and to ignore the flanking words. On the second display, a single string of letters was presented and participants had to indicate whether the string formed a legal word or not (lexical decision). Fox found that lexical decision responses to L2 words were slowed down if these target words were semantically related to L1 words that had been presented as flankers on the previous display (i.e., participants needed more time to indicate that SEL [salt] was a French word when the word *pepper* had been used as flankers on the previous display). Negative priming was also observed from L2 flanking words on L1 targets if both words were translation equivalents (i.e., *sel* used as a flanker and SALT as the target).

Second, primed lexical decision tasks have shown that the recognition of a target word is facilitated when it is preceded by a tachistoscopically presented prime which is a semantic associate in the other language. This

[3] In agreement with the literature on bilingualism, I will use term 'lexical representations' with reference to orthographic representations. Entries in the phonological mental lexicon will be labelled as 'phonological representations'.

effect is about 75% the size of the effect observed when the semantic associate is in the same language (Francis, 1999). For instance, de Groot and Nas (1991) found that for Dutch-English bilinguals, lexical decision responses to the word GIRL were faster not only after the prime *boy* but also after the prime *jongen* (the Dutch word for *boy*). Similarly, Grainger and Frenck-Mestre (1998) observed that English-French bilinguals were faster to decide that the letter sequence TREE formed a legal English word when it followed the French translation prime *arbre* than when it followed the unrelated prime *balle* [ball]. This effect was found despite the fact that primes were presented only for 43 ms and could not be recognized by the participants. The translation priming effect was reliably stronger when participants were asked to perform a semantic categorization task instead of a lexical decision task, yielding further evidence that the origin of the effect was semantical (for a discussion of semantic vs. lexical-associative interpretations of translation priming, I refer to the *General Discussion* section of Chapter 4).

Third, semantic comparisons (e.g. semantic categorization tasks) between words from different languages have been shown to take no longer than comparisons between words of the same language, again suggesting the integration of semantic information between languages (Potter, So, Voneckardt, & Feldman, 1984; see Francis, 1999, for a review).

Fourth, Dijkstra and colleagues (Dijkstra, Grainger, & Van Heuven, 1999; Dijkstra, Van Jaarsveld, & Ten Brinke, 1998) found that lexical decisions are faster for cognates than for interlingual homographs and language-unique words of the same frequency. Cognates are translation equivalents which have also orthographic and/or phonological overlap (e.g. *apple* - *appel* in English and Dutch). Interlingual homographs also share orthography and phonology (partially) but not meaning (e.g. *room* is a word both in English and Dutch, but means *cream* in Dutch). Language unique words are words that only exist in one language. The faster reaction times to the cognates than to the two other types of words can only be explained if one accepts meaning similarity in L1 and L2. Such a facilitatory effect should not occur if semantic representations (at least for cognates) are not shared across languages.

Finally, using fMRI, Illes et al. (1999) measured the brain activity of proficient bilinguals performing a semantic categorization task (abstract vs. concrete words) in L1 and L2. These authors were unable to find significant differences in brain activity between both language conditions. In both L2 and L1, there was enhanced activation in the left inferior prefrontal cortex, which is in line with findings from previous monolingual studies.

In summary, there is a large consensus that the semantic representations of translation equivalents are at least partially shared across languages. For a more detailed discussion of this topic, I refer to Kroll (1993) and Kroll and de Groot (1997).

THE ORGANIZATION OF THE ORTHOGRAPHIC LEXICON IN BILINGUALS

Because equivalent words in different languages usually have different forms (except for cognates; see above), the intuitively most appealing theory about the lexical organization of a bilingual person is that there are two different orthographic lexicons ('mental dictionaries'): one for L1 and one for L2. In addition, it seems to make sense that if a person is reading in one language, only the lexicon of this language is active and the other is temporarily inhibited. As indicated by Dijkstra, Grainger and Van Heuven (1999), this is a model with language-dependent storage and language selective access. However, there is increasing evidence that these intuitively plausible ideas are wrong. In the present section, I will present some of the recent findings on this issue. For a more thorough overview, I refer to Dijkstra and Van Heuven (2002) and Brysbaert (1998).

First, de Groot, Delmaar and Lupker (2000) showed that the processing of interlingual homographs (see earlier, e.g. *room*) in a translation recognition task ('are both words translation equivalents?' yes/no) was slower compared to the processing of matched control words. This was especially the case when the homograph reading to be selected was the less frequent of the two homograph's readings. This is a strong demonstration of lexical influences from a non-active language.

Secondly, Dijkstra, Timmermans and Schriefers (2000) showed that Dutch-English bilinguals respond slower to interlingual homographs (e.g. *room*) than to words which only exist in L2 in a lexical go/no-go task (press a button only if the target is a word in L2). This strongly suggests that L1 lexical representations are active to a certain degree when performing a task in L2 which does not explicitly require activation of L1 knowledge (unlike the translation recognition paradigm of de Groot et al., 2000, discussed above).

Third, Van Heuven, Dijkstra and Grainger (1998) manipulated the number of orthographically similar words in L1 and L2. An English word like *left*, for example, has quite some English neighbors that differ from the word in only one letter position (e.g. *deft, heft, lift, loft, lent, lest*); it also has many Dutch neighbors of this type (e.g. *heft, lift, lest, leut*). Other words have few neighbors both in English and in Dutch (e.g. *deny*), few neighbors in English but many in Dutch (e.g. *keen*), or many neighbors in English but few in Dutch (e.g. *coin*). Previous monolingual research has indicated that word recognition depends on the neighborhood size of the word: The more orthographically similar words a target word has, the easier it is to process the word. Van Heuven et al. presented the above four types of English words to native Dutch speakers and found that reaction times not only depended on the number of orthographic neighbors in English but also on the number of orthographic neighbors in Dutch. Again, this indicates that L1 (Dutch) word forms were activated during L2 (English) word recognition, even though L1 was irrelevant for the task.

Fourth, Van Hell and Dijkstra (2002) recently showed that L2 and even L3 lexical knowledge also influences L1 lexical access in an exclusive native language context. They reported faster lexical decision responses of Dutch – English – French trilinguals for L1 targets having L2 and L3 near-cognate (i.e. orthographically nearly identical) translation equivalents (e.g. *banaan – banana - banane*) than for control words. This shows that L2 (and even L3) lexical representations are accessed during L1 word recognition and that their activation is strong enough to influence L1 representations.

The previous findings offer strong evidence against language-selective lexical access. This also calls into question the language-dependent storage assumption (i.e., that L1 and L2 words are represented in different lexicons), but does not rule out this possibility. Stronger neuropsychological evidence for separate lexicons would be provided if a double dissociation were reported between a bilingual patient who was dyslexic in L1 but not in L2 (provided there are no obvious differences between the languages; e.g. that both make use of the same alphabet) and another patient who (for the same language pair) was dyslexic in L2 but not in L1. However, such a dissociation has not yet been reported, and will probably never be reported. Also, similar findings from the aphasia literature, such as the observation that the ability to speak is sometimes affected differently in L1 and L2 (e.g. Fabbro, 2001b), can not be considered irrefutable evidence, because the observed partial dissociation may be due to differently affected control mechanisms. Several authors have suggested that, when a specific language is not available anymore, it is not necessarily because its neural substrates have been physically destroyed, but maybe because its control system has been weakened, in terms of increased inhibition, raised activation threshold, or unbalanced distribution of resources among the various languages (Green, 1986; Paradis, 1998). For a more detailed discussion of this topic, I refer to Fabbro (2001a).

In summary, there is a growing body of evidence that lexical access during visual word recognition is not language-specific (whether both lexicons are stored separately or not). L1 lexical representations are activated to a certain extent during L2 word processing and vice versa. For a more detailed discussion of this topic, I refer to Dijkstra and Van Heuven (2002).

THE ORGANIZATION OF THE PHONOLOGICAL LEXICON IN BILINGUALS

Whereas studies on bilingualism have almost exclusively focused on semantic and orthographic/lexical representations, models of monolingual word recognition illustrate that words are represented through at least one more level, i.e. through phonological representations (entries in a phonological lexicon). Surprisingly, only a few bilingual studies (e.g. Jared & Kroll, 2001; Van Wijnendaele & Brysbaert, 2002) were directly aimed at investigating whether the early activation of phonological representations during word recognition is language-independent (just as for lexical representations) or not.

First, based on the claim that visual word recognition implies automatic, pre-lexical phonological coding (see Frost, 1998, for a review), Brysbaert, Van Dyck and Van de Poel (1999) reasoned that it is very likely that an automatic (not strategically controlled) grapheme-to-phoneme conversion (GPC) occurs for all grapheme-phoneme correspondences mastered by bilinguals. This conversion should take place before a language selection mechanism (if any, see earlier) gets involved in the word recognition process. To investigate this, the authors started from the finding in monolingual studies that the recognition of target words (e.g. MAIL) is facilitated by the presentation of a masked homophonic prime (e.g. *male*) relative to a graphemic control prime (e.g. *mall*). Then, using Dutch-French bilinguals, they showed that the same effect can be found for the recognition of L2 targets using cross-lingual (L1) homophone primes. For instance, recognition of the target word OUI [yes] was facilitated by the prime *wie* [who]. Because the L1 prime is only a homophone of the L2 target according to L1 GPC rules, this strongly suggests that L1 phonological representations are accessed during L2 visual word recognition.

Second, Jared and Kroll (2001) found further evidence for this claim using a naming task with French-English bilinguals. In this study, L2 words which have word-body enemies in L1 (e.g. the English word *bait* contains the word body *ait* which is pronounced differently in French) were named slower than control words (e.g. the word *bump* contains the letter sequence *ump* which is illegal in French).

Third, similar results were obtained by Dijkstra et al. (1999). In this study, Dutch-English participants completed a lexical decision task and a progressive demasking task with English (L2) words varying in their degree of semantic, orthographic and phonological overlap. Importantly for the present section, responses to L2 words having phonological overlap with existing L1 words were slower than to control words, even though L1 phonology was not relevant for the task at hand.

Fourth, Van Wijnendaele and Brysbaert (2002) replicated the cross-language phonological priming effect of Brysbaert et al. (1999), using the same stimuli, but reversing the language dominance of the participants (French-Dutch bilinguals). This offers strong evidence for the claim that letter strings are automatically and pre-lexically coded through L2 GPC rules[4], even when performing a task in L1.

[4] Note that the use of the term 'GPC rules' does not imply that grapheme-to-phoneme conversion follows strict, mutually exclusive 'rules', according to which one 'rule' eventually activates one phonological representation. Rather, the studies discussed in this section suggest that graphemes and word forms access phonological representations through parallel activation (across languages), much in the way orthographic representations are activated in interactive activation models (see below).

Fifth, Jared and Kroll (2001) also replicated their study described above with English-French bilinguals, and found that L1 words with L2 word-body enemies were indeed named slower than control words, similar to the effects found for L2 words having L1 word-body enemies. However, this effect was only present when the participants had just named a block of L2 filler words.

Finally, these findings from the visual word recognition literature are not without analogues in auditory word recognition. For example, Weber and Cutler (2004) examined lexical competition in non-native spoken word recognition. They observed that Dutch-English bilinguals hearing English (L2) target words (e.g. *desk*) made longer eye fixations on distractor pictures with Dutch (L1) names phonologically related to the English target (e.g. a picture of a *deksel* [lid]). Contrastingly, Dutch listeners showed no activation of the English words (picture of a *desk*, given the target word *deksel*). These results suggest that L1 phonological representations are activated during L2 input, and that competition is greater for L2 than for L1 listeners. Similarly findings were reported by Marian, Spivey and Hirsch (2003). They found that Russian-English bilinguals instructed in English to "*pick up the marker*" often looked at a stamp, because its Russian translation equivalent (*marka*) has phonological overlap with the English spoken word.

In conclusion, these few studies on language-independent activation of phonological representations strongly suggest that visually presented words are always automatically processed through L1 GPC rules, even when reading in L2. Evidence for the opposite statement is mixed. Whereas the results of Van Wijnendaele and Brysbaert (2002) clearly show that L2 phonological representations are accessed during L1 word recognition, findings of Jared and Kroll (2001) suggest that the activation in these representations may only be strong enough to influence L1 processing if L2 GPC rules have recently been active. For a more comprehensive review of studies on this topic, I refer to the introductory section of Chapter 4.

MODELS OF BILINGUALISM

Before going into details about the various experiments included in this dissertation, I will briefly present the three most influential models of bilingual language organization. As pointed out in the third and final part of this chapter, each of these models will appear to be especially important for one of the following chapters of this dissertation.

THE REVISED HIERARCHICAL MODEL: LEXICO-SEMANTIC LINKS

The Revised Hierarchical Model (RHM) of Kroll and colleagues (e.g. Kroll & Stewart, 1994; Kroll & de Groot, 1997) has probably been the most influential model of bilingualism during the last ten years. It was mainly used to explain how lexical and semantic representations interact when words are translated from L1 to L2 (forward translation) and vice versa (backward translation). The model does not include phonological representations. It includes two language-specific lexical stores[5] and a common semantic system (for a figure, see Chapter 2, Figure 1).

Unlike earlier models (e.g. Weinreich, 1953), the lexical and semantic components of the RHM are fully interconnected. However, the strength of these connections varies and is asymmetrical. The links between L1 word forms and the meaning they represent are assumed to be stronger than those between L2 word forms and their semantic representations. Similarly, the lexical word-word connections are thought to be stronger from L2 to L1 than the other way around. This is because L2 words are initially learned by associating them with L1 translations. As a consequence, the model predicts that forward translation is more likely to engage semantic mediation than

---

[5] Note that the storage-dependent assumption is not an essential element of the model. The model would make the same predictions with respect to the translation process if the L1 and L2 lexical items were part of a single, combined lexicon, in which translations had direct lexical connections.

backward translation, certainly during the first stages of L2 language acquisition. However, L2 lexico-semantic links are assumed to strengthen as L2 proficiency increases and words have been encountered in many meaningful contexts ('the developmental hypothesis'). In highly L2 proficient bilinguals, these connections are assumed to become strong enough to influence backward translation.

There is a large body of evidence supporting the asymmetry and developmental assumptions of the RHM. For instance, Sholl, Sankaranarayanan and Kroll (1995) found that forward translation was significantly enhanced when the involved concepts had been primed by the presentation of pictures in a first phase of the experiment. The priming effect was not present in the backward translation condition, suggesting a less conceptually mediated translation process. Similarly, Kroll and Stewart (1994) manipulated the semantic relatedness of the stimuli within a word list for a translation task: half of the lists contained words from a single semantic category, half contained words from different categories. This manipulation of context did not affect word naming and backward translation, but it did have an effect on forward translation. It was more difficult to translate words in the blocked lists (presumably due to increased competition at the semantic level) than in the mixed lists. Furthermore, backward translation was faster than forward translation, in line with the strong and direct lexical connections postulated from L2 words to L1 words. Similar effects have been reported by other authors (Cheung & Chen, 1998; Fox, 1996; Keatley, Spinks, & Degelder, 1994).

As for the developmental hypothesis of the RHM, Talamas, Kroll and Dufour (1999) found greater interference of semantically related false translations in a translation recognition task when participants were highly proficient in L2, whereas less proficient bilinguals suffered more from form-related distractors. This suggests that the latter group of participants relied more on lexical information for the translation task. Other evidence was reported by Kroll, Michael, Tokowicz and Dufour (2002) who observed that the speed advantage for translation of cognates (relative to non-cognates) is particularly strong in beginning bilinguals. Again, this suggests that the overlap in the lexical representations between translation equivalents is more important in the early stages of second language acquisition that in later stages. For a more comprehensive overview of studies supporting the RHM, I refer to Kroll and de Groot (1997).

THE DISTRIBUTED FEATURE MODEL: SEMANTICS

Whereas the RHM makes clear and very explicit assumptions and predictions about the nature and strength of L1 and L2 lexico-semantic connections, the semantic and lexical stores are actually black boxes. This is not the case for the Distributed Feature Model of de Groot and colleagues (see Figure 1, e.g. de Groot, 1992; de Groot, Dannenburg, & van Hell, 1994; van Hell & de Groot, 1998b), which has mainly focused on the organization of semantic representations. Although few researchers still doubt that multilinguals have a single semantic system accessed by all the languages known (see earlier), this does not imply that the meaning of all words in the different languages must be exactly the same. Indeed, bilinguals often have the feeling that a word (or expression) in one language does not have a translation equivalent with exactly the same meaning. To describe the relative overlap in meaning representations for different types of concepts, word meanings are represented as sets of distributed features in the model. The overlap in meaning, indexed by the number of shared features, depends on the type of word being represented. One of the main tenets of the distributed feature model is that concrete words (e.g. *ball*) have more similar meanings (indicated by a larger feature overlap) across languages than abstract words (e.g. *justice*). This is based on the assumption that the functions of the objects to which concrete words refer will be similar across languages and cultures. In contrast, because abstract words tend to be used in different contexts across languages, they will be less similar in meaning. Evidence for this theory comes from de Groot (1992) for example, who observed faster translation production and recognition responses to concrete control words than to matched abstract words. For further evidence, I refer to Kroll and de Groot (1997). Note that Van Hell and de Groot (1998a) later adapted the distributed feature model to account for different degrees of overlap at both the lexical and conceptual level.

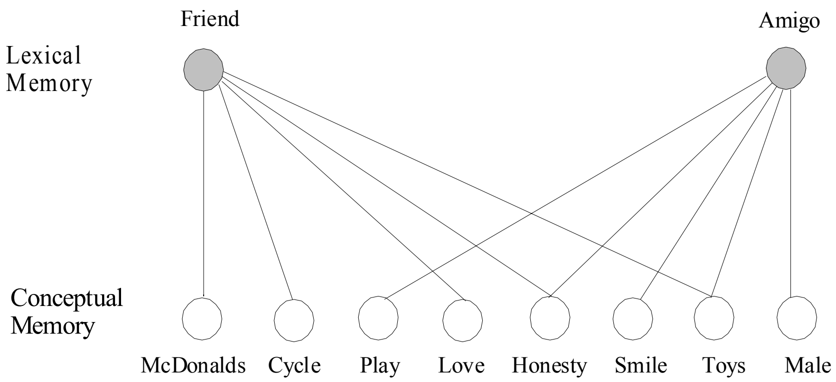


Figure 1: the Distributed Feature Model (English - Spanish Example)

THE BILINGUAL INTERACTIVE ACTIVATION MODEL: ORTHOGRAPHIC LEXICAL REPRESENTATIONS

Whereas the RHM is primarily concerned with the strength of lexico-semantic connections and the distributed feature model is focused on semantic representations (with a large role for word concreteness), the Bilingual Interactive Activation (BIA) model deals with bilingual orthographic/lexical organization. BIA is a very powerful, implemented computational model which has been able to explain a large body of findings in bilingual visual word recognition, without recurrence to the intuitively appealing -but wrong-idea of language-specific lexicons (for a review, see Dijkstra & Van Heuven, 2002). The model is an extension of the Interactive Activation model for monolingual word recognition (e.g. McClelland & Rumelhart, 1981),

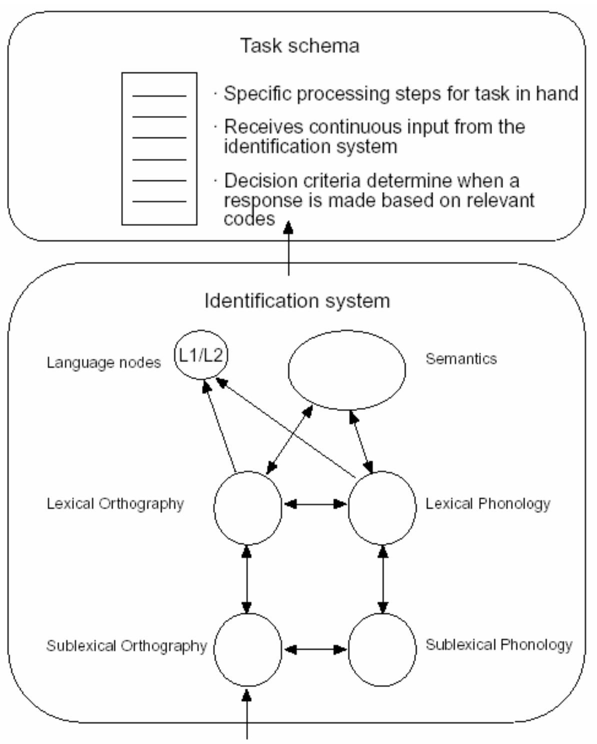


Figure 2: the BIA+ Model

containing language, word, letter and feature nodes. In the model, all L2 and L1 words are represented into a unitary word-level system. Lexical access during word recognition is initially non-selective, as word activation is affected by competing items from both languages. For evidence supporting the model, I refer to the studies discussed earlier in the section on bilingual lexical organization. Note that the BIA model has recently been updated (now the BIA+ model, see Dijkstra & Van Heuven, 2002). This model (depicted in Figure 2) does not contain any top-down connections from language nodes to word nodes, unlike the earlier BIA model. Effects of non-linguistic (e.g. stimulus list composition) and linguistic (e.g. sentence) context, which have shown to influence word recognition (e.g. Dijkstra et al., 2000), are dealt with by a task-decision system, which only receives input from the (fundamentally language non-selective) word identification system.

A MODEL OF BILINGUAL PHONOLOGICAL REPRESENTATIONS?

As noted earlier, research on language-selective functioning of the bilingual language processing system has mainly focused on lexical and semantic representations. As a consequence, there is no model of bilingual phonological processing at present. Still, as can be seen in Figure 2, the sketch of the recent BIA+ model already contains phonological (and semantic) representations. However, these have not been implemented yet and are still black boxes. From the research on bilingual phonological processing (see earlier), it will be clear that any future model will have to be structurally language non-selective with regard to the activation of phonological representations, much in the way the BIA(+) model operates with orthographic/lexical representations.

THE PRESENT DISSERTATION

In the current dissertation, I will present several studies which are aimed at gaining knowledge with respect to one of the above-mentioned representational levels of language and respective models of bilingualism. Following the same order as above, the three following parts of this dissertation deal, respectively, with L2 lexico-semantic organization, semantic representations and phonological representations.

FIRST PART: L2 LEXICO-SEMANTIC CONNECTIONS

In the first part (Chapter 2) of this dissertation, the primary focus is on L2 lexico-semantic connections, i.e. the interaction between L2 lexical and semantic representations. As discussed earlier, the only model in the literature which makes very explicit assumptions and predictions about this issue is the RHM of Kroll and colleagues (e.g. Kroll & Stewart, 1994; Kroll & de Groot, 1997). According to the RHM, L2 word forms (in contrast to L1 word forms) have weaker connections than L1 words with the semantic representations that they represent. Therefore, it is assumed that L2 word forms primarily access the meaning system through lexical connections with their L1 translation equivalents. The strength of these connections is only believed to increase notably in high levels of L2 proficiency, and is not assumed to differ for different types of words. One prediction that follows from this asymmetrical lexico-semantic organization is that backward translation (from L2 to L1) is not semantically mediated, as opposed to forward translation (from L1 to L2).

Whereas there is a large body of evidence supporting these claims (see earlier), I believe there are reasons, both of an empirical and theoretical nature, to review several aspects of the model. First, a small number of empirical studies have recently reported strong semantic effects in backward translation, which is incompatible with the asymmetry hypothesis of the RHM (e.g. Duyck & Brysbaert, 2002; Bloem & La Heij, 2003; La Heij, Hooglander, Kerling, & Van der Velden, 1996). Secondly, as a theoretical argument, because the asymmetry hypothesis of the RHM states that L2 lexico-semantic connections are much weaker than for L1, the model assumes strong facilitatory lexical links between L2 and L1 translation equivalents for backward translation. This may be hard to reconcile with the mechanism of lateral inhibition between lexical representations which is present in many interactive activation models (such as the BIA model, see earlier) (T. Dijkstra, personal communication, October 14^{th}, 2003).

Starting from the previous observations, we decided to investigate the nature of L2 lexico-semantic organization and the assumptions of the RHM with respect to this issue in more detail. To this end, we designed a series of experiments in which we tried to find indications of the existence of (strong) L2 lexico-semantic links. Because number words are probably the words with the most confined semantic representations, and therefore maybe the strongest lexico-semantic connections, we investigated form-to-meaning mappings of L2 number words. Using number words as stimuli also has the advantage that semantic involvement can often be easily determined by looking at effects of their primary meaning, i.e. their size (e.g. the number magnitude effect: small numbers are processed faster than large numbers). We also investigated whether the nature of lexico-semantic connections of L2 number words interacts with L2 proficiency, as a test of the developmental hypothesis of the RHM. These experiments are presented in Chapter 2.

SECOND PART: SEMANTIC REPRESENTATIONS: CONCRETENESS

Whereas the first part of this dissertation deals with L2 lexico-semantic organization, the primary focus in the second part (Chapter 3) is on semantic representations. As discussed earlier, in the distributed feature model of de Groot and colleagues (see Figure 1, e.g. de Groot, 1992; de Groot et al., 1994; van Hell & de Groot, 1998b), it is assumed that word meanings are represented as sets of distributed features. The overlap in meaning between translation equivalents, indexed by the number of shared features, depends on the type of word being represented. One of the main tenets of the model is that concrete words (e.g. *ball*) have more similar meanings (indicated by a larger feature overlap) across languages than abstract words (e.g. *justice*). This assumption is supported by several empirical observations, such as the faster translation production and recognition responses for concrete words than for abstract words (de Groot, 1992).

In Chapter 3, further evidence was sought for the importance of the word variable word concreteness (or imageability). More specifically, we investigated whether new word forms are mapped onto existing visual (semantic) information when they are first acquired, provided such information is available. Such a finding would offer further evidence (see also Part 1) that new word forms are not only associated with existing lexical information (i.e. the word forms of the translation equivalent), but are also linked to existing stored information of a non-lexical nature. Moreover, because the information in the presented experiments is of a visual nature, it would offer support for the role of word concreteness as postulated by the distributed feature model. If new word forms are mapped onto existing visual information, this suggests the existence of certain lexico-semantic links for concrete, but not for abstract words. Because we wanted to design a very strong test of semantic coding during word learning, we used associative word learning, which is a method that certainly tempts to learn words by forming lexical word associations. Coding of the visual information was investigated by means of a selective interference paradigm, which is widely used in working memory research. In Chapter 7, I will discuss similar own findings that were recently obtained (not included in the present dissertation) with respect to the word variable concreteness (or imageability), using a more typical psycholinguistic paradigm (i.e. masked priming in a lexical decision task).

THIRD PART: ACCESS TO PHONOLOGICAL REPRESENTATIONS

In the third and final part of this dissertation, I will focus on the issue of language-selective access to phonological representations in bilinguals. As indicated earlier, research on bilingualism has mainly focused on lexico-semantic organization (Chapter 2), semantic representations (Chapter 3) and language selective access to orthographic/lexical representations (e.g. the BIA model, see earlier). Consequently, both empirical and modeling work on phonological coding in bilinguals is scarce. The few studies which have dealt with this issue suggest that phonological coding may be language independent (see earlier), much in the way that access to orthographic lexical representations is language-independent. For instance, Van Wijnendaele and Brysbaert (2002) showed that the recognition of L1 words (e.g. OUI [yes]) in French-Dutch bilinguals is facilitated by L2 homophone primes (e.g. *wie* [who]). This strongly suggests that words are coded through L2 GPC rules, even when performing a task in L1. In the final part of this dissertation, we tried to find further (stronger) evidence for this claim, by investigating whether the recognition of L1 (Dutch) and L2 (English) target words can be facilitated by related L2 and L1 pseudohomophones (i.e. nonwords which sound like a L2 or L1 word related to the target's translation equivalent). Such an effect should only occur if cross-lingual phonological coding is sufficiently strong and fast enough to even activate associated words during the first stages of word recognition. Therefore, this would offer evidence for strong language-independent interactions of orthographic/lexical, semantic and phonological representations in bilinguals (see also Brysbaert, Van Wijnendaele, & Duyck, 2002). Six experiments dealing with this issue are reported in Chapter 4. In Chapter 5, I will report a study which was designed to investigate whether the degree of language-independent phonological coding interacts with L2 proficiency. This was motivated by a claim made by Gollan, Forster and Frost (1997), who hypothesized that reliance on phonology decreases with increasing L2 proficiency. Therefore, this may have an effect on the size of cross-lingual phonological effects like those observed by Van Wijnendaele and Brysbaert (2002). The findings from these two chapters provide the necessary evidence for the primary assumptions which will have to be made when creating a future model of bilingual phonological representations (regarding the interdependence of the phonological lexicons for example). Finally, in the last chapter of this third part (Chapter 6), I will present WordGen, a tool which substantially simplifies the generation of word and nonword stimuli for psycholinguistic (priming) experiments. This tool was extensively used for the generation of the stimuli used in Chapter 4. After this third empirical part, I bring this dissertation to a close with Chapter 7, in which I present the theoretical implications and general conclusions which can be drawn from the previous chapters.

REFERENCES

Bamgbose, A. (1994). Nigeria's choice. *UNESCO Courier,* 22-31.

Bloem, I., & La Heij, W. (2003). Semantic Facilitation and Semantic Interference in Word Translation: Implications for Models of Lexical Access in Language Production. *Journal of Memory and Language, 48*(3), 468-488.

Brysbaert, M. (1998). Word Recognition in Bilinguals: Evidence Against the Existence of Two Separate Lexicons. *Psychologica Belgica, 38*(3-4), 163-175.

Brysbaert, M., Van Dyck, G., & Van de Poel, M. (1999). Visual Word Recognition in Bilinguals: Evidence From Masked Phonological Priming. *Journal of Experimental Psychology: Human Perception and Performance, 25*(1), 137-148.

Brysbaert, M., Van Wijnendaele, I., & Duyck, W. (2002). On the temporal delay assumption and the impact of non-linguistic context effects. *Bilingualism: Language and Cognition, 5*(3), 199-201.

Cheung, H., & Chen, H. C. (1998). Lexical and Conceptual Processing in Chinese-English Bilinguals: Further Evidence for Asymmetry. *Memory & Cognition, 26*(5), 1002-1013.

De Bot, K. (1992). A Bilingual Production Model: Levelt's Speaking Model Adapted. *Applied Linguistics, 13*, 1-24.

de Groot, A. M. B. (1992). Determinants of Word Translation. *Journal of Experimental Psychology: Learning, Memory and Cognition, 18*(5), 1001-1018.

de Groot, A. M. B., Dannenburg, L., & van Hell, J. G. (1994). Forward and Backward Word Translation by Bilinguals. *Journal of Memory and Language, 33*(5), 600-629.

de Groot, A. M. B., Delmaar, P., & Lupker, S. J. (2000). The Processing of Interlexical Homographs in Translation Recognition and Lexical Decision: Support for Non-Selective Access to Bilingual Memory. *Quarterly Journal of Experimental Psychology, 53A*(2), 397-428.

de Groot, A. M. B., & Nas, G. L. J. (1991). Lexical Representation of Cognates and Noncognates in Compound Bilinguals. *Journal of Memory and Language, 30*(1), 90-123.

Dijkstra, A., Grainger, J., & Van Heuven, W. J. B. (1999). Recognition of Cognates and Interlingual Homographs: the Neglected Role of Phonology. *Journal of Memory and Language, 41*(4), 496-518.

Dijkstra, A., & Van Heuven, W. (2002). The Architecture of the Bilingual Word Recognition System: From Identification to Decision. *Bilingualism: Language and Cognition, 5*(3), 175-197.

Dijkstra, A., Van Jaarsveld, H., & Ten Brinke, S. (1998). Interlingual Homograph Recognition: Effects of Task Demands and Language Intermixing. *Bilingualism: Language and Cognition, 1*(1), 51-66.

Dijkstra, T., Timmermans, M., & Schriefers, H. (2000). On Being Blinded by Your Other Language: Effects of Task Demands on Interlingual Homograph Recognition. *Journal of Memory and Language, 42*(4), 445-464.

Duyck, W., & Brysbaert, M. (2002). What Number Translation Studies Can Teach Us About the Lexico-Semantic Organisation in Bilinguals. *Psychologica Belgica, 42*(3), 151-175.

Fabbro, F. (1999). The Neurolinguistics of Bilingualism: an Introduction. Psychology Press.

Fabbro, F. (2001a). The Bilingual Brain: Cerebral Representation of Languages. *Brain and Language, 79*, 211-222.

Fabbro, F. (2001b). The Bilingual Brain: Bilingual Aphasia. *Brain and Language, 79*, 201-210.

Fox, E. (1996). Cross-Language Priming From Ignored Words: Evidence for a Common Representational System in Bilinguals. *Journal of Memory and Language, 35*(3), 353-370.

Francis, W. S. (1999). Cognitive Integration of Language and Memory in Bilinguals: Semantic Representation. *Psychological Bulletin, 125*, 193-222.

Frost, R. (1998). Toward a Strong Phonological Theory of Visual Word Recognition: True Issues and False Trails. *Psychological Bulletin, 123*(1), 71-99.

Gollan, T. H., Forster, K. I., & Frost, R. (1997). Translation Priming With Different Scripts: Masked Priming With Cognates and Noncognates in Hebrew-English Bilinguals. *Journal of Experimental PsychologyLearning, Memory and Cognition, 23*(5), 1122-1139.

Grainger, J., & Frenck-Mestre, C. (1998). Masked Priming by Translation Equivalents in Proficient Bilinguals. *Language and Cognitive Processes, 13*(6), 601-623.

Green, D. (1986). Control, Activation, and Resource: a Framework and a Model for the Contol of Speech in Bilinguals. *Brain and Language, 27*, 210-223.

Grosjean, F. (1982). *Life with Two Languages: An Introduction to Bilingualism*. Cambridge, MA: Harvard University Press.

Grosjean, F. (1992). Another View of Bilingualism. In Harris, R. (Ed.), *Cognitive Processing in Bilinguals.* (pp. 51-62). Amsterdam: Elsevier.

Hakuta, K. (1986). Mirror of Language: the Debate on Bilingualism. New York: Basic Books.

Hakuta, K., & Diaz, R. (1985). The Relationship Between Bilingualism and Cognitive Ability: A Critical Discussion and some New Longitudinal Data. In Nelson, K. E. (Ed.), *Handbook of Psycholinguistic Research.* (pp. pp. 1-32). San Diego: Academic Press.

Illes, J., Francis, W. S., Desmond, J. E., Gabrieli, J. D. E., Gover, G. H., Poldrack, R., Lee, C. J., & Wagner, A. D. (1999). Convergent Cortical Representation of Semantic Processing in Bilinguals. *Brain and Language, 70*, 347-363.

Jared, D., & Kroll, J. F. (2001). Do Bilinguals Activate Phonological Representations in One or Both of Their Languages When Naming Words? *Journal of Memory and Language, 44*(1), 2-31.

Keatley, C. W., Spinks, J. A., & Degelder, B. (1994). Asymmetrical Cross-Language Priming Effects. *Memory & Cognition, 22* (1), 70-84.

Kroll, J. F. (1993). Assessing Conceptual Representations for Words in a Second Language. In Schreuder, R. and Weltens, B. (Eds.), *The Bilingual Lexicon.* (pp. 53-82). Amsterdam: John Benjamins.

Kroll, J. F., & de Groot, A. M. B. (1997). Lexical And Conceptual Memory in the Bilingual: Mapping Form to Meaning in Two Languages. In de Groot, A. M. B. and Kroll, J. F. (Eds.), *Tutorials In Bilingualism: Psycholinguistic Perspectives.* (pp. 201-224). Mahwah, NJ: Lawrence Erlbaum Publishers.

Kroll, J. F., Michael, E., Tokowicz, N., & Dufour, R. (2002). The Development of Lexical Fluency in a Second Language. *Second Language Research, 18*, 137-171.

Kroll, J. F., & Stewart, E. (1994). Category Interference in Translation and Picture Naming - Evidence for Asymmetric Connections Between Bilingual Memory Representations. *Journal of Memory and Language, 33*(2), 149-174.

La Heij, W., Hooglander, A., Kerling, R., & Van der Velden, E. (1996). Nonverbal Context Effects in Forward and Backward Word Translation: Evidence for Concept Mediation. *Journal of Memory and Language, 35*(5), 648-665.

Landry, R. G. (1974). A Comparison of Second Language Learners and Monolinguals on Divergent Thinking Tasks at the Elementary School Level. *Modern Language Journal, 58*, 10-15.

Levelt, W.J.M. (1989). *Speaking: from Intention to Articulation* . Cambridge, Massachusetts: Bradford.

Marian, V., Spivey, M., & Hirsch, J. (2003). Shared and Separate Systems in Bilingual Language Processing: Converging Evidence From Eyetracking and Brain Imaging. *Brain and Language, 86*(1), 70-82.

McClelland, J. L., & Rumelhart, D. E. (1981). An Interactive Activation Model of Context Effects in Letter Perception : An Account of Basic Findings. *Psychological Review, 88*(5), 375-407.

Paradis, M. (1998). Language and Communication in Multilinguals. In Stemmer, B. and Whitaker, H. (Eds.), *Handbook of Neurolinguistics.* (pp. San Diego, CA: Academic Press.

Potter, M. C., So, K. F., Voneckardt, B., & Feldman, L. B. (1984). Lexical and Conceptual Representation in Beginning and Proficient Bilinguals. *Journal of Verbal Learning and Verbal Behavior, 23*(1), 23-38.

Sholl, A., Sankaranarayanan, A., & Kroll, J. F. (1995). Transfer Between Picture Naming and Translation - a Test of Asymmetries in Bilingual Memory. *Psychological Science, 6*(1), 45-49.

Talamas, A., Kroll, J. F., & Dufour R. (1999). From Form to Meaning: Stages in the Acquisition of Second-Language Vocabulary. *Bilingualism: Language and Cognition, 2*(1), 45-58.

van Hell, J. G., & de Groot, A. M. B. (1998a). Conceptual Representation in Bilingual Memory: Effects of Concreteness and Cognate Status in Word Association. *Bilingualism: Language and Cognition, 1*, 193-211.

van Hell, J. G., & de Groot, A. M. B. (1998b). Disentangling Context Availability and Concreteness in Lexical Decision and Word Translation. *Quarterly Journal of Experimental Psychology, 51A*, 41-63.

Van Hell, J. G., & Dijkstra, T. (2002). Foreign Language Knowledge Can Influence Native Language Performance in Exclusively Native Contexts. *Psychonomic Bulletin & Review, 9*(4), 780-789.

Van Heuven, W. J. B., Dijkstra, A., & Grainger, J. (1998). Orthographic Neighborhood Effects in Bilingual Word Recognition. *Journal of Memory and Language, 39*(3), 458-483.

Van Wijnendaele, I., & Brysbaert, M. (2002). Visual Word Recognition in Bilinguals: Phonological Priming from the Second to the First Language. *Journal of Experimental Psychology: Human Perception and Performance, 3*, 619-627.

Weber, A., & Cutler, A. (2004). Lexical Competition in Non-Native Spoken-Word Recognition. *Journal of Memory and Language, 50*(1), 1-25.

Weinreich, U. (1953). *Languages in Contact*. New York: The Linguistic Circle Of New York.

CHAPTER 2
FORWARD AND BACKWARD NUMBER TRANSLATION REQUIRES CONCEPTUAL MEDIATION BOTH IN BALANCED AND UNBALANCED BILINGUALS

***Revised manuscript submitted for publication*[1, 2]**

It is much debated whether translation is semantically mediated or based on word–word associations at the lexical level. In two experiments with Dutch (L1) – French (L2) bilinguals, we showed that there is a semantic number magnitude effect in both forward and backward translation of number words: it takes longer to translate number words representing large quantities (e.g. *acht*, *huit* [eight]) than small quantities (e.g. *twee*, *deux* [two]). In a third experiment, we replicated these effects with number words that had been acquired only just before the translation task. Finally, it was shown that our findings were not due to the restricted semantic context of our stimuli. These findings strongly suggest that translation processes can be semantically mediated in both directions, even at low levels of L2 proficiency.

---

[1] This paper was co-authored by Marc Brysbaert

[2] The authors are indebted to Judith Kroll, Robert Remez, Michael Thomas and an anonymous reviewer for excellent comments on an earlier version of this paper. We would also like to thank Jolien De Brauwer for carrying out a replication of Experiment 3.

INTRODUCTION

Bilingual people are able not only to understand messages in two different languages, but also to translate information from their first language (L1) to their second (L2) (forward translation) and vice versa (backward translation). An important question with respect to this ability is whether word translation relies on direct word – word associations in the lexicon, or whether it requires activation of the meaning of the words. The aim of the present study was threefold: (a) to replicate semantic effects during forward translation, which almost all translation studies have reported (e.g. Kroll & Stewart, 1994; Sholl, Sankaranarayanan, & Kroll, 1995; Cheung & Chen, 1998; La Heij, Hooglander, Kerling, & Vandervelden, 1996; Sánchez-Casas, Davis, & Garcia-Albea, 1992), , (b) to search for indications of semantic mediation in backward translation and (c) to investigate whether the nature of the translation process depends on additional factors, such as L2 proficiency and semantic context. We will report four experiments in which we addressed these questions by looking for a semantic number magnitude effect in a number naming and translation task. The dominant view in the literature on this issue is provided by Kroll and Stewart's Revised Hierarchical Model (RHM) of bilingual memory (Kroll & Stewart, 1994; Kroll & de Groot, 1997), which we describe in some detail below.

CONCEPTUAL MEDIATION IN WORD TRANSLATION

In a series of papers, Kroll and colleagues (Kroll & Stewart, 1994; Kroll & de Groot, 1997) developed a model of bilingual memory (see Figure 1) that can explain a broad range of findings (for reviews see Kroll, 1993; Kroll & de Groot, 1997). First, in the model a distinction is made between lexical representations (representing word forms) and semantic representations (representing word meanings)(for a recent review of evidence concerning this assumption, see Smith, 1997). Second, it is assumed that the conceptual representations are shared among the languages, whereas the lexical representations are language-specific. So, there are two lexicons (one for L1, and one for L2), connected to a unitary semantic system. Finally, the connections between the different parts in the model have two interesting features: They are asymmetric and their importance changes in the course of second language acquisition (the developmental hypothesis).

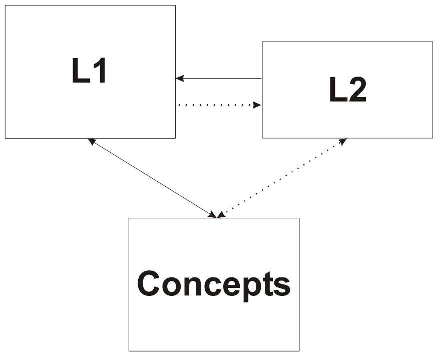


Figure 1. The RHM of bilingual memory (Kroll & de Groot, 1997). Solid lines represent stronger links than dotted lines.

The asymmetry hypothesis states that the links between the L1 lexicon and the semantic system are stronger than the links between the L2 lexicon and the semantic system, because word meanings are more strongly activated by L1 words than by L2 words. In contrast, the direct word – word connections at the lexical level are stronger from L2 to L1 than the other way around. The reason for this is that L2 words are often learned by associating them with their L1 translations. Because of the asymmetries in the connections, forward translation is more likely to engage conceptual activation than backward translation. Backward translation in turn depends more on direct lexical connections. Support for this asymmetry assumption was reported by Sholl et al. (1995). On forward translation trials, they found facilitation when the concepts had been primed by the presentation of pictures in a first phase of the experiment. The priming effect was not present in the backward translation condition, suggesting a less conceptually mediated translation process. The asymmetries of the connections are believed to decline as L2 proficiency increases. This is the developmental aspect of the model. Support for this hypothesis was reported by Talamas, Kroll, and Dufour (1999), who found greater interference of semantically related false translations in a translation recognition task when the participants were highly proficient in L2, whereas less proficient bilinguals suffered more interference from form-related words. For a more detailed review of the findings supporting the different assumptions of the RHM, we refer to Kroll and de Groot (Kroll & de Groot, 1997; see also Kroll & Stewart, 1994; Sánchez-Casas et al., 1992; Cheung & Chen, 1998; Kroll, Michael, Tokowicz, & Dufour, 2002).

In the experiments described below, we tested to what extent these assumptions and following predictions of the RHM hold in a number word translation task. Before presenting these experiments, we briefly review what is currently known about number processing (for a more detailed review, see Brysbaert, in press).

SEMANTIC ACTIVATION IN NUMBER PROCESSING

Numbers provide a very appealing set of stimuli to study translation processes, as bilinguals have three sets of symbols to represent the same concept: Arabic digits (e.g. '*3*'), L1 number words, (e.g. '*drie*' in Dutch), and L2 number words (e.g. '*trois*' in French). This makes it possible not only to study translation from L1 to L2 and from L2 to L1, but also from a common logographic symbol to either L1 or L2. Moreover, as will be shown below, activation of the underlying conceptual information depends on the symbol format and the nature of the task. In addition, Arabic digits do not take much longer to name than number words, contrary to the naming of object pictures (Ferrand, 1999).

The meaning of a number primarily refers to the magnitude represented by the number. These magnitudes can be understood quite well with the metaphor of a number line (e.g., Brysbaert, 1995; Dehaene, 1992). All integers (from 1 to at least 15) form an ordered continuum oriented from left (small) to right (large). So, when the magnitudes of two numbers have to be compared, this is easier when the distance between the numbers is large (e.g., to indicate the larger number of the pair 2-8) than when it is small (e.g., to indicate the larger number of the pair 7-8, Moyer & Landauer, 1967).

Similarly, the processing of a number is primed when immediately before a number with a close magnitude has been presented than when a number with a more distant magnitude has been presented (Reynvoet & Brysbaert, 1999). In addition, participants react faster with their left hand to small numbers and with their right hand to large numbers than vice versa (the so-called SNARC effect, e.g., Dehaene, Bossini, & Giraux, 1993).

Further evidence indicates that the magnitude information is activated more rapidly for small numbers than for large numbers. So, it is easier to select the larger digit of the number pair 2-3 than of the number pair 7-8 (Moyer & Landauer, 1967). Using eye movement registrations, Brysbaert (1995, Experiment 1) found that this effect is not entirely due to the comparison process, but also to the encoding speed of the numbers. In this experiment, participants had to read three Arabic numerals going from 0 to 99, and decide whether the middle number fell in-between the two outer numbers (e.g. 23 41 65) or not (e.g. 23 65 41). The most important variable to predict the reading time of the first numeral turned out to be the (logarithm of the) number magnitude. More importantly, in a subsequent experiment (Brysbaert, 1995, Experiment 2), the same magnitude effect was replicated when participants simply had to read the three numerals, and indicate whether a fourth, new numeral was one of the first three numbers or not. Because the fourth numeral could not be seen until all three numerals of the initial set had been read, this effect could not be due to the comparison process, but rather to the encoding of the numbers.

There are several lines of evidence indicating that the number magnitude effect is semantic in origin (i.e., related to the meaning of the numbers) and not due to associations between lexical representations. For a start, the distance related priming effect (see earlier) is the same within-notations as across-notations. Thus, the target word *six* is primed as much by the prime *5* as by the prime *five* (Reynvoet, Brysbaert, & Fias, 2002). Similarly, the same distance effect in number comparison is found with digits, number words, and even random sets of dots as stimulus materials (e.g. Dehaene, Dupoux,

& Mehler, 1990; Buckley & Gillman, 1974; Foltz, Poltrock, & Potts, 1984). Furthermore, the quantity priming effect is symmetrical (e.g. the target *6* is primed equally well by the prime *5* than by *7*), and not asymmetric (with stronger priming in the forward prime-target direction) as an associative hypothesis would predict (e.g. Koechlin, Naccache, Block, & Dehaene, 1999; Duyck & Brysbaert, 2002). Also, Dehaene and Akhavein (1995) observed effects of numerical distance when the lexical distance between items (according to theories of number processing) was kept constant. Finally, brain imaging studies have shown that in all number comparison tasks, irrespective of the input format, a region in the parietal cortex is active, which is not active in non-numerical word processing tasks but which is active in other analog magnitude estimation tasks (Fias, Lammertyn, Reynvoet, Dupont, & Orban, 2003; Naccache & Dehaene, 2001; Pinel, Dehaene, Riviere, & Lebihan, 2001; Pesenti, Thioux, Seron, & De Volder, 2000). For further evidence for the semantic origin of these effects, we refer to Koechlin, Naccache, Block and Dehaene (1999)

More importantly for the present study, magnitude information is not required for all number processing tasks. There is quite some evidence that the processing of number words is only semantically mediated if the experimental task requires the activation of certain semantic information. For instance, Fias, Reynvoet, and Brysbaert (2001; see also, Fias, 2001) showed that the word *five* was read equally fast when it was displayed together with the digit *3* as when it was presented together with the digit *5*. In contrast, responses to the word *five* were faster when the word was presented together with the digit *5* than when it was presented with the digit *3* in a parity judgment (odd/even) task. The finding that number words can be named without semantic mediation is in line with most models of visual word recognition, which assume the existence of non-semantic routes for word naming. The situation is less clear for Arabic input, with authors claiming that digits can be named without semantic mediation (e.g., Campbell, 1994; Cipolotti & Butterworth, 1995; Dehaene, 1992), whereas others reject this possibility (e.g., Brysbaert, 1995, see earlier; Fias, 2001; McCloskey, 1992).

Both groups of authors agree, however, that number magnitude is more rapidly activated from Arabic input than from verbal input, as can be concluded from the finding that participants find it more difficult to select the physically larger number in the number pair 2 – 8 than in the number pair 2 – 8, whereas no such size congruency effect is observed in the pairs two – eight vs. two – eight (e.g., Ito & Hatta, 2003).

EXPERIMENT 1

From the previous section, it will be clear that number translation provides an interesting new paradigm to test Kroll and Stewart's (1994) RHM. Given that word naming does not explicitly require access to semantic information, L1 and L2 naming of respectively L1 and L2 number words is not expected to show a magnitude effect. The Dutch word *twee* (two) is not expected to be named faster than the word *acht* (eight). The same applies to the corresponding French number words *deux* and *huit*. It is less clear whether the naming of Arabic numbers will involve a magnitude effect: On the basis of rapid activation of magnitude information from Arabic input, one might expect to find such an effect both in L1 and in L2 naming of digits. So, the naming of the Arabic digit *2* in Dutch and French could be faster that the naming of *8.* Most importantly for the present study, however, the presence or absence of a semantic magnitude effect in the translation of number words allows us to directly test the asymmetry hypothesis of the RHM. Because number words have been shown to activate their underlying semantic information in certain (semantic) tasks (see earlier), and because forward translation implies conceptual mediation, a magnitude effect should be found when Dutch-French bilinguals translate Dutch (L1) number words into French (L2). Hence, the word *twee* (two) should be translated faster into French than the number word *acht* (eight). In contrast, no semantic number magnitude effect is expected for these bilinguals when French (L2) number words have to be translated into Dutch (L1), as this task is more likely to occur through word – word associations at the lexical level, and therefore does not require access to semantic representations. So, for Dutch-French bilinguals the French number word *deux* (two) should be translated into Dutch as fast as *huit* (eight), except maybe for very proficient bilinguals (see the developmental hypothesis of the RHM).

To test these predictions of the RHM, we designed an experiment in which Arabic numbers and both L1 (Dutch) and L2 (French) number words had to be named in both L1 and L2. We also manipulated L2 proficiency to check for interactions with possible number magnitude effects. In addition, a short delayed naming task was administered after the actual experiment. In this task, participants were asked to delay responses for more than a second, so that semantic processing of the stimulus was finished before the response had to be given. This allows us to control for differences in voice key sensitivity to the response onset and other theoretically irrelevant variables which could confound the number magnitude effect that is of interest in this study.

METHOD

Participants. Twenty-two first-year university students participated for course requirements. All of them were native Dutch speakers, and mainly used this language in everyday life. Eleven of them had started learning French at school between 10 and 13 years of age. This group will be referred to as ‘unbalanced’ bilinguals. They did not study their L2 at university, had no L2 speaking relatives and did not practice their L2 on any other regular basis. The other eleven students had been raised in a Dutch-French bilingual setting from birth, were practically equally proficient in both languages (but indicated Dutch as L1), and will therefore be referred to as the ‘balanced bilinguals’.

Materials. All stimuli were presented on a standard 15” VGA color monitor, as yellow characters on a black background. Stimulus presentation was computer driven by a PC equipped with a voice key which was connected through the gameport. All Arabic digits, Dutch and French number words representing quantities from 1 to 12 were selected as stimuli.

Design. The experiment had a 2 (L2 Proficiency: ‘unbalanced’ versus ‘balanced’) x 2 (Naming Language: L1 versus L2) x 3 (Stimulus format: Arabic numbers, L1 number words and L2 number words) x 12 (Number Magnitude) full factorial design. Except for L2 proficiency, all variables were manipulated within-subjects.

Procedure. All participants completed two blocks (L1 naming and L2 naming) of 360 trials. The order of these blocks was counterbalanced with L2 proficiency. Within each block, 10 series of 36 randomly ordered trials were presented, corresponding to every number magnitude from 1 to 12 in each of the three stimulus formats (Arabic, L1, L2). Hence, the participants did not know which stimulus format would appear before the beginning of each trial. Only naming language was blocked. Each trial started with the presentation of a fixation stimulus (‘*’; the plus-sign was not used as a fixation point because of its mathematical meaning) for 500 ms. After that time, the stimulus was replaced by the target, which remained visible until pronunciation of the target triggered the voice key. The Inter Trial Interval (ITI) was 1000 ms. The experiment lasted for about 50 minutes, including a little break.

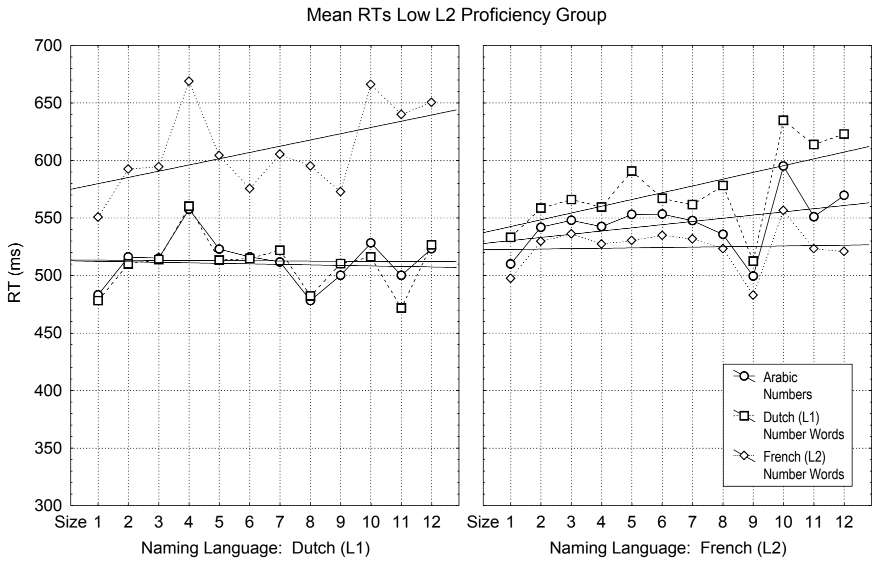
As noted earlier, we also ran a delayed number word naming task after the actual experiment in order to control for theoretically irrelevant variables such as voice key onset sensitivity. In this task, 12 participants (unbalanced bilinguals) were asked to delay responses for more than a second. All participants completed eight blocks of 24 trials. Within each block, all Dutch and French number words representing magnitudes from 1 to 12 were presented in a random order. In each trial, the target was presented centred on the screen for 1000 ms, followed by a black screen for another 1000 ms. Then, a question mark was presented, indicating that the participant had to name the target word seen before as fast as possible. The ITI was 1000 ms.

As soon as there was doubt concerning the accuracy of time registration (e.g. due to irrelevant noise), the trial was excluded from the data.

RESULTS

Variance Analysis. The proportion of invalid trials in the immediate naming experiment due to naming errors or faulty time registration was 6.1%. These trials were excluded from all analyses. An ANOVA was performed with L2 Proficiency as between-subject variable and Naming Language, Stimulus Format, and Number Magnitude as repeated measures factors. The dependent variable was the mean RT across correct trials. Mean RTs for both L2 proficiency groups as a function of Naming Language, Stimulus Format and Number Magnitude are presented in Figure 2. The backward translation condition can be found in the left part of the figures, whereas forward translation is plotted in the right part of the figures.

The main effect of L2 Proficiency reached significance, $F(1, 20) = 3.26$, $MSE = 160709.2$, $p < .05$ (one-tailed). Mean RTs for balanced and unbalanced bilinguals were respectively 510 ms and 546 ms. As the two



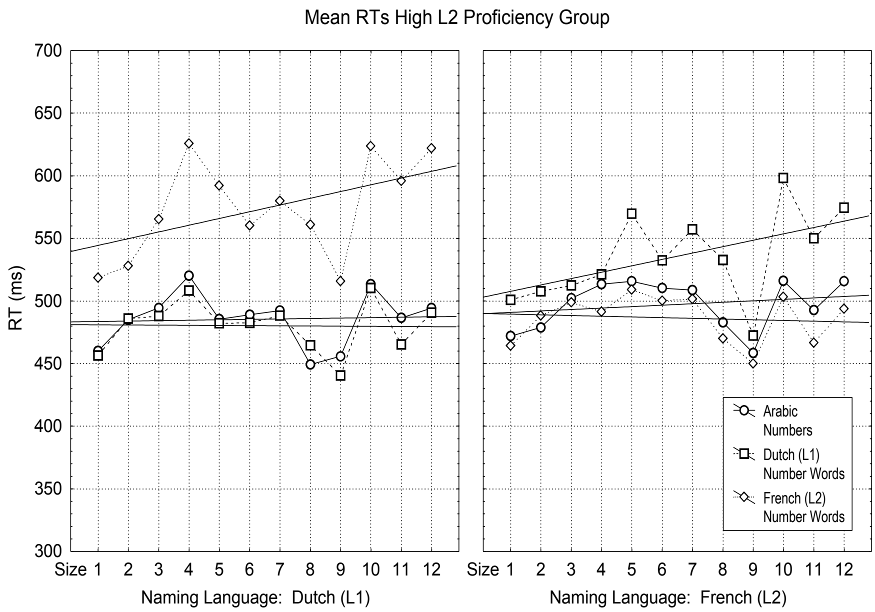


Figure 2: Mean naming RTs by naming language, stimulus format and number magnitude (Experiment 1), plotted separately for both proficiency groups. Straight lines represent best linear fit according to a least squares criterion.

almost identical graphs in Figure 2 suggest, Proficiency did not interact with any other factor in the design. The effect of Naming Language was not significant, $F < 1$. Naming in Dutch took 529 ms, while French naming took 527 ms. Unbalanced bilinguals were slightly slower for French (L2) naming (M = 548 ms) than for Dutch (L1) naming (M = 544 ms), but this difference was not significant, $F < 1$. Balanced bilinguals showed a tendency towards the reverse pattern [L2 naming (M = 506 ms) was slightly faster than L1 naming (M = 513 ms)], but again this difference was not significant, $F < 1$. A Post hoc comparison using Tukey's HSD test showed that backward translation was significantly slower (M = 592 ms) than forward translation (M = 555 ms), $p < .001$, as opposed to the predictions based on the RHM. This difference was significant both for unbalanced (M = 610 ms vs. M = 575 ms) and balanced bilinguals (M = 574 ms vs. M = 536 ms), ps < .01.

The main effects of Stimulus Format ($F(2, 40) = 76.12$, $MSE = 2596.4$, $p < .001$) and Number Magnitude ($F(11, 220) = 27.55$, $MSE = 2352.6$, $p < .001$)

were significant, but these effects were embedded in an important three way interaction with naming language, $F(22, 440) = 11.86$, $MSE = 705.6$, $p < .001$. Indeed, as can be seen in Figure 2, the effect of Number Magnitude appears to be present in only some of the Stimulus Format x Naming Language conditions. These effects of Number Magnitude will be investigated in more detail by means of regression analyses in the next section.

Regression Analysis. *T*o assess the importance of number magnitude independent of number frequency[3] and the delay to activate the voice key, regression analyses were performed according to the procedure for repeated measures data described by Lorch and Myers (1990, Method 3), with number magnitude, number word frequency and mean delayed naming RTs as predictors.

The regression weights for the six conditions [i.e. 2 naming language (L1 versus L2) x 3 stimulus formats (Arabic, L1, and L2)] are displayed in Table 1. Most importantly, the regression weights of number magnitude differed significantly from zero in both the forward and the backward translation condition, respectively $t(21) = 5.135$, $p < .001$ (p values for two-tailed testing) and $t(21) = 7.940$, $p < .001$ (for a detailed statistical explanation of the computational procedure of these tests, see Lorch & Myers, 1990). These regression weights did not differ significantly from each other, $t(21) = 1.322$, $p > .20$. Hence, conceptual mediation was not larger in forward than in backward translation.

[3] We wanted to control for frequency effects because Gielen, Brysbaert, and Dhondt (1991) found a significant correlation between number magnitude and number frequency ($r = -.621$, $p < .01$). So, because smaller numbers are more frequent and are thus processed faster, it is possible that any effect of number magnitude in the data is confounded by effects of number frequency. Therefore, we included the number word frequency measures as reported by Gielen et al. (1991) in our analysis. Note that Dehaene and Mehler (1992) showed that the frequencies of numbers are very similar in different languages.

Naming Language	Stimulus Format		Intercept	Number Magnitude	Delayed Naming RT	Frequency
L1 Naming (Dutch)	Arabic Numbers	Y =	220	+ 0.22 NM	+ 0.70 D**	+ 0.03 F
	L1 Number Words (Dutch)	Y =	183	+ 0.00 NM	+ 0.77 D**	+ 0.05 F
	L2 Number Words (French)	Y =	80	+ 5.08 NM**	+ 1.26 D**	- 0.01F
L2 Naming (French)	Arabic Numbers	Y =	231	+ 2.71 NM**	+ 0.75 D**	- 0.04 F
	L1 Number Words (Dutch)	Y =	215	+ 6.94 NM**	+ 0.73 D**	+ 0.04 F
	L2 Number Words (French)	Y =	235	+ 0.71 NM	+ 0.72 D**	- 0.02 F

Table 1. The regression equations for the six naming language x stimulus format conditions (Experiment 1) according to the procedure described by Lorch and Myers (1990) (** $p < .01$).

The size of the number magnitude effect did not differ between unbalanced and balanced bilinguals. This was the case for both forward and backward translation ($t < 1$). To further increase the power of our analysis; we also tested whether the size of the magnitude effect correlated with the difference in mean RTs between L1 and L2 naming (as a measure of L2 proficiency)[4]. Consistent with the previous results, this correlation was very weak and not significant for either direction of translation (backward translation: $r = -.090$, $p > .69$; forward translation: $r = -.126$, $p > .57$), Thus, the magnitude effect in number word translation did not interact with L2 proficiency.

The number magnitude effect was not significant for within-language number word naming (L1 – Dutch: $t < 1$; L2 - French: $t(21) = 1.218$, $p > .23$), nor for L1 (Dutch) naming of Arabic digits, $t < 1$. In contrast, the regression weight of number magnitude differed significantly from zero for

[4] We subtracted mean RTs for the size 2 condition from mean RTs for size 8 as a measure of the number magnitude effect. Note that different measures (e.g. the regression weights of number magnitude in the regression analyses) led to very similar results. We thank Michael Thomas for this suggestion.

the L2 (French) naming of Arabic digits, $t(21) = 3.631$, $p < .01$. The regression weights of Frequency never reached significance.

All regression weights of the Delayed Naming RTs were significant. This confirms that the RTs were indeed influenced by sensitivity differences in the triggering of the voice key by the different number names. Related studies in speech production have also acknowledged this problem. Jescheniak and Levelt (1994) for example dealt with it by subtracting delayed response RTs from non-delayed RTs, and using the resulting values as the dependent variable in the analysis. If we follow this approach (Figure 3), we see that taking into account the delayed naming RTs does not affect the number magnitude effects observed in forward and backward translation. Indeed, regression analyses with the dependent variable RT immediate naming minus delayed naming yielded virtually identical results as the ones mentioned above.

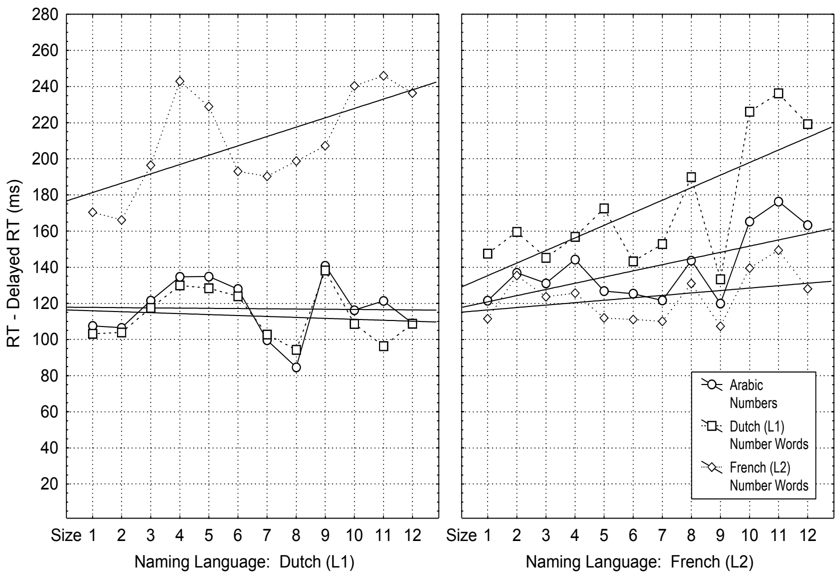


Figure 3. Mean naming RTs minus delayed naming RTs by naming language, stimulus format and number magnitude (Experiment 1). Straight lines represent best linear fit according to a least squares criterion.

DISCUSSION

The results of Experiment 1 are quite clear. As predicted by the RHM (Kroll & Stewart, 1994), we obtained a reliable effect of number magnitude on forward translation. It took longer to translate L1 number words representing large quantities (e.g. *acht* [eight]) than number words representing small quantities (e.g. *twee* [eight]). This strongly suggests conceptual mediation, since magnitude information is not stored at the lexical level. However, contrary to the predictions of the RHM, such an effect was also obtained in backward translation: Translation was slower for large L2 number words (e.g. *huit* [eight]) than for small L2 number words (e.g. *deux* [eight]). Moreover, the effect was equally strong for both directions of translation and did not interact with L2 proficiency. Hence, it seems that translation was conceptually mediated in both directions to the same degree for both balanced and less proficient bilinguals. Another finding which is not consistent with the predictions of the model concerns the speed of the translation processes: Backward translation was found to be significantly slower than forward translation.

An account of the observed number magnitude effects in terms of the correlated predictor word frequency (see Footnote 1) cannot easily explain why the effect only emerges in the translation conditions, and not in the naming conditions. Indeed, the regression analyses confirmed that the magnitude effects found were not due to effects of number word frequency. Frequency did not have an effect in any of the conditions. As the frequency effect is usually situated at the lexical level, this is further indirect evidence that the translations were not based on direct word – word associations. For a more detailed theoretical discussion of these results, we refer to the *General Discussion*.

As expected, we did not find a semantic effect for number word naming, contrary to number word translation, since this is not a semantic task. In contrast, we did find a magnitude effect in L2 naming of Arabic numbers.

This is not surprising since processing of Arabic numbers seems to trigger fast conceptual activation (see earlier). However, we did not obtain an effect for L1 naming of Arabic numbers, contrary to Brysbaert (1995). The cause of the absence of such an effect could lie within the restricted range of digits used: whereas Brysbaert (1995) presented numbers from 0 to 99, we only used digits from 1 to 12. An inspection of Brysbaert's data shows that the logarithmic number magnitude effect was not very clear for numbers smaller than 10. We believe this finding only adds further importance to the effects found in translation trials. Finally, the effects of the delayed naming predictor showed the importance of controlling for the speed with which a voice key is activated by different sequences of sounds.

EXPERIMENT 2

Although the results described above seem quite straightforward, one might object that the random presentation procedure used in Experiment 1 rendered the input stimulus format (but not the output language) unpredictable. It is possible that this introduced switching costs between trials (e.g. when an L1 target is followed by an L2 target). For example, Meuter and Allport (1999) showed in an Arabic number naming task that switching naming language resulted in a time cost on the trial following the switch. The cost was larger when switching from L2 to L1 than in the reverse condition. Similar switching costs may have occurred in Experiment 1, even though there are theoretical reasons to believe that it is unlikely that such costs have influenced the obtained magnitude effects (see the *Discussion* section). This alternative switching cost account cannot be ruled out by a reanalysis of our data, as inclusion of the previous trial stimulus format as a variable leads to very few observations per cell (on average 3.33, if all trials were correct). Moreover, the random algorithm which steered stimulus presentation did not guarantee a minimum of one observation per cell. Therefore, we decided to repeat Experiment 1 with a procedure that blocked not only naming language, but also stimulus format.

METHOD

Participants. Twelve first-year university students participated for course credit. They all belonged to the group of bilinguals labelled before as 'unbalanced'.

Materials. All stimulus materials were identical to Experiment 1. The software used was adapted to make stimulus presentation blocked by stimulus format.

Design. The experiment had a 2 (Naming Language: L1 versus L2) x 3 (Stimulus format: Arabic numbers, L1 number words and L2 number words) x 12 (Number Magnitude) full factorial design. All variables were manipulated within-subjects. L2 proficiency was not manipulated because it did not interact with any variable in Experiment 1.

Procedure. All participants completed two series of three blocks, each consisting of 120 trials. Naming language was Dutch (L1) in one series and French (L2) in the other. The three blocks within a series corresponded to the three stimulus formats (Arabic numbers, L1 number words, L2 number words). Within each block, all numbers from 1 to 12 were presented 10 times. Order of trials within blocks, blocks within series, and series was determined at random. The procedure for a trial was identical to that used in Experiment 1.

RESULTS

Variance Analysis. The proportion of invalid trials due to naming errors or faulty time registration was 7.0%. These trials were excluded from all analyses. An ANOVA was performed with Naming Language, Stimulus Format, and Number Magnitude as repeated measures factors. The dependent variable was the mean RT across correct trials. Mean RTs as a function of Naming Language, Stimulus Format and Number Magnitude are presented in Figure 4. The backward translation condition can be found in the upper left part of the figure, while forward translation corresponds to the upper graph in the right part of the figure.

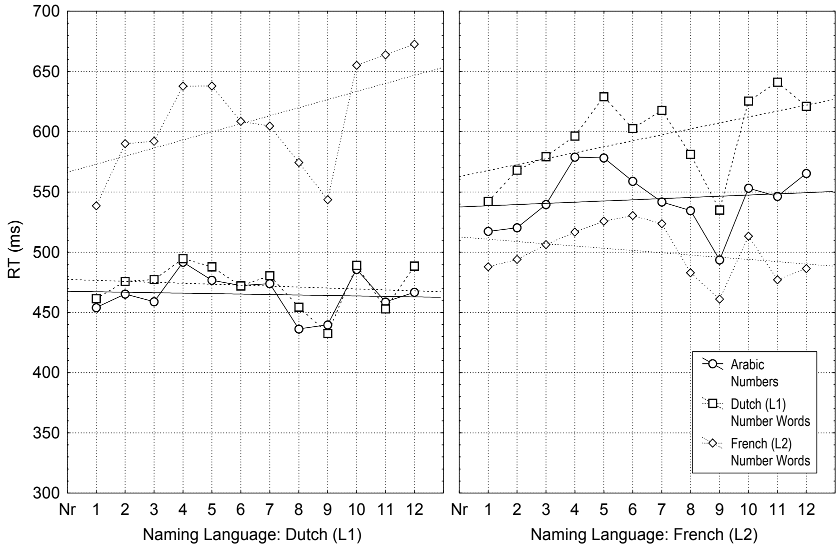


Figure 4. Mean naming RTs by naming language, stimulus format and number magnitude (Experiment 2 – blocked presentation). Straight lines represent best linear fit according to a least squares criterion.

The effect of Naming Language was significant, $F(1, 11) = 7.71$, $MSE = 26484.9$, $p < .018$. Respective means for Dutch and French naming were 516 and 547 ms. Forward translation ($M = 595$ ms) tended to be faster than backward translation ($M = 610$ ms), but Tukey's HSD test showed that this difference was not significant, $p < .693$. The main effects of Stimulus Format ($F(2, 22) = 25.86$, $MSE = 7212.9$, $p < .001$) and Number Magnitude ($F(11, 121) = 24.91$, $MSE = 1568.1$, $p < .001$) were significant, but these effects were embedded in a three way interaction with naming language, $F(22, 242) = 10.67$, $MSE = 726.8$, $p < .001$. As can be seen in Figure 4, and similar to Experiment 1, the effect of Number Magnitude was only present in some of the Stimulus Format x Naming Language conditions. Again, these effects will be analyzed in detail in the following regression analyses.

Regression Analysis. The regression analyses were again performed by the procedure for repeated measures data described by Lorch and Myers (1990, Method 3). Regression weights for the six Naming Language (L1 versus L2) x Stimulus Format (Arabic, L1, and L2) conditions are displayed in Table 2. Similar to Experiment 1, the regression weights of Number Magnitude differed significantly from zero in both forward and backward translation conditions, respectively $t(11) = 2.718$, $p < .020$ and $t(11) = 4.949$, $p < .001$. These regression weights did not differ significantly from each other ($t < 1$), although the effect of number magnitude tended to be somewhat larger in the backward translation condition. Note that an effect of Frequency was found in the forward translation condition, while this effect was not present in backward translation.

Naming Language	Stimulus Format		Intercept	Number Magnitude	Delayed Naming RT	Frequency
L1 Naming (Dutch)	Arabic Numbers	Y =	217	+ 0.12 NM	+ 0.58 D**	+ 0.07 F
	L1 Number Words (Dutch)	Y =	171	- 0.43 NM	+ 0.74 D**	+ 0.06 F
	L2 Number Words (French)	Y =	164	+ 5.18 NM**	+ 1.22 D**	- 0.14 F
L2 Naming (French)	Arabic Numbers	Y =	322	+ 0.02 NM	+ 0.81 D**	- 0.23 F*
	L1 Number Words (Dutch)	Y =	330	+ 3.88 NM*	+ 0.88 D**	- 0.25 F*
	L2 Number Words (French)	Y =	155	- 1.10 NM	+ 1.00 D**	- 0.07 F

Table 2. The regression equations for the six naming language x stimulus format conditions (Experiment 2) according to the procedure described by Lorch and Myers (1990) (* $p < .05$; ** $p < .01$).

A comparison of Tables 1 and 2 strongly suggests that the blocking of the stimuli made no difference for the number magnitude effect. This was confirmed by a statistical analysis (Lorch & Myers, 1990). There was no difference at all between the number magnitude regression weights for backward translation (Experiment 1: $B = 5.08$; Experiment 3: $B = 5.18$; $t < 1$). For forward translation, the regression weights were slightly higher in Experiment 1 (Experiment 1: $B = 6.94$; Experiment 2: $B = 3.88$), but this difference was not significant ($t(32) = 1.449$, $p > .15$). Also, the blocked presentation procedure had no effect on overall mean RTs (Experiment 1: $M = 528$ ms; Experiment 2: $M = 531$ ms; $F < 1$). If only RTs from the unbalanced bilinguals who participated in Experiment 1 were taken into account, mean RTs for Experiment 2 (only unbalanced participants) were slightly faster, but this difference again was far from significant (Experiment 1: $M = 546$ ms; Experiment 2: $M = 531$ ms; $F < 1$).

DISCUSSION

The present experiment shows that the findings of Experiment 1 were not caused by switching costs due to the random presentation of different stimulus formats (e.g. Meuter & Allport, 1999). When stimulus format was blocked, exactly the same effects of number magnitude were found in both forward and backward translation. It took longer to translate L1 and L2 number words representing larger quantities than small quantities. The fact that mixed and blocked stimulus presentation yielded the same results, is not inconsistent with the literature on switching costs, as previous studies on bilingualism have reported switching costs when the output language changed (e.g. Meuter & Allport, 1999), but not when the input language changed (e.g. Thomas & Allport, 2000). In addition, there is quite some evidence that the initial stages of visual word recognition in bilinguals are not language-specific, and consequently not subject to switching costs (Nas, 1983; Altenberg & Cairns, 1983; Von Studnitz & Green, 1997; Dijkstra, Grainger, & Van Heuven, 1999; Dijkstra, Timmermans, & Schriefers, 2000; Brysbaert, Van Dyck, & Van de Poel, 1999; Dijkstra & Van Heuven, 2002; Van Wijnendaele & Brysbaert, 2002). More importantly, the fact that equivalent number magnitude effects were found for forward and backward translation in two different studies with different stimulus presentations, adds credit to our claim that number translation is conceptually mediated in both translation directions. We will go into further details about the theoretical implications of this finding in the *General Discussion*.

EXPERIMENT 3

In the previous experiments, we found evidence for an equivalent number magnitude effect in forward and backward number translation, both for unbalanced and balanced bilinguals. As these findings are not entirely compatible with the predictions of the RHM (stronger semantic mediation in forward translation than in backward translation; more asymmetry at low levels of L2 proficiency), we decided to explore more in depth the limits of our findings and to eliminate some possible confounds in our stimulus materials.

A criticism against Experiments 1 and 2 might be that the unbalanced bilinguals were already too proficient. Indeed, mean RTs were reasonably fast for a translation task. This was probably partly due to the fact that the number words are among the first acquired and most frequent L2 words. Also, because of political reasons, Belgian high school students receive rather extensive L2 teaching. Hence, it might be that the number magnitude effect in backward translation was a manifestation of a very high L2 proficiency overall in our participant population, in accordance with the developmental hypothesis of the RHM. Therefore, it would be interesting to investigate how the number magnitude translation effect manifests itself at much lower levels of L2 proficiency.

As there is little conformity in the literature regarding the assessment of L2 proficiency, we decided to experimentally manipulate this variable, rather than making use of indirect measures of L2 proficiency. Therefore, we designed a learning experiment in which participants learned the first 15 number words from a so-called foreign language (‘Estonian’). In reality, these number words were non-existing words, in order to exclude any inherent structure in the stimuli and any correspondences with other languages known to the participants (e.g., the Estonian words for 1-15 are *üks*, *kaks*, *kolm*, *neli*, *viis*, kuus, *seitse*, *kaheksa*, *üheksa*, *kümme*, *üksteist*, *kaksteist*, *kolmteist*, *neliteist* and *viisteist*). Immediately after the participants had acquired these words, a number word translation task similar to the previous experiments was administered. If the number magnitude translation effect does not occur in such an experiment, this would offer support for the developmental hypothesis of the hierarchical model. In contrast, if the magnitude effect manifests itself even though these words were only acquired several minutes earlier, this would offer substantial evidence in favor of the hypothesis that number word forms are mapped onto the number line from the very first encounters of these words.

Using a new number language also allowed us to better control our stimulus materials. As indicated above, we took out any inherent structure from the number "words" (as a matter of fact, each participant received a different number-word mapping). In addition, we controlled the number of times participants came across the various L2 words. Although it is common practice in L2 language acquisition to teach all number words up to 12 together, it is still possible that bilinguals in later reading more often encountered some numbers than others, and that this accounted for (some of) our magnitude effect.

METHOD

Participants. Twenty first-year university students participated for course requirements. They all belonged to the group of bilinguals labelled before as 'unbalanced'.

Materials. Fifteen legal six-letter non-words that followed the Dutch orthographic rules (*soltil*; *fidara*; *lacron*; *nelima*; *sipron*; *kodrim*; *sertir*; *badiks*; *dreksa*; *kummer*; *dorter*; *fistar*; *gabiro*; *meltir*; *pridar*) were created and randomly assigned as the translation equivalents of one of the Dutch number words between *een* (one) and *vijftien* (fifteen). Each participant got a different, random mapping of numbers and words. Participants were told that the stimuli were Estonian number words. We made use of non-words, because this allowed us to control for word length (six letters), prior experience with the stimuli, and any correlation between the number words and the magnitudes they stand for (e.g., in real Estonian there is a correlation of .93 between the length of the number words and the magnitudes they represent for the integers 1-15). We included number words ranging from 1 to 15 in order to slightly extend the magnitude range from the previous experiments.

Design. Similar to Experiment 1 and 2, the experiment had a 2 (Naming Language: L1 versus L2) x 3 (Stimulus format: Arabic numbers, L1 number words and L2 number words) x 15 (Number Magnitude) full factorial design. All variables were manipulated within-subjects.

Procedure. The experiment consisted of four parts: a learning phase, a test phase, an experimental phase and a delayed naming task.

The Learning Phase. During each trial, a Dutch number word was presented for six seconds together with its 'Estonian' translation. The participants were instructed to memorize these number words so that they would be able to recall a number word given its translation equivalent. No hints were given concerning possible memorization strategies. Each trial was preceded by a fixation stimulus for 500 ms. Number word combinations were presented in a random order. The following test phase started when each number word combination was shown once.

The Test Phase. This phase aimed at measuring memorization performance after each learning phase. During each trial, a Dutch or 'Estonian' number word was presented on the screen. The participants were instructed to enter the respective translation of the word by means of the keyboard within a ten second timeframe. A recall test was chosen instead of a recognition test in order to avoid learning during the test phase as much as possible. If more than 85% of all trials were correct (i.e. 13 out of 15), participants proceeded to the experimental phase, after a little break. If not, the learning phase was administered again. On average, participants needed 10 (range 6-16) exposures to the learning phase before they reached the criterion.

The Experimental Phase. This phase was identical to the number word translation task of Experiment 1. The only difference concerned the amount of number magnitude conditions (i.e. fifteen instead of twelve, see earlier).

The Delayed Naming Task. The day after these phases, the participants completed a delayed naming task of Dutch and 'Estonian' words, identical to that of Experiment 1. This made it possible to include the same predictors in our regression model as in the previous analyses.

RESULTS

Variance Analysis. The proportion of invalid trials due to naming errors, memorization failure or faulty time registration was 25.3%. This is quite high because we opted to use an 85% criterion (and not a 100% criterion, see earlier) during the word learning test phase in order to avoid overlearning and to keep L2 proficiency as low as possible. Of course, this caused a number of false trials in addition to the other errors which are generally committed in this type of experiments. The invalid trials were excluded from all analyses. An ANOVA was performed with Naming Language, Stimulus Format, and Number Magnitude as repeated measures factors. The dependent variable was the mean RT across correct trials (estimating the resulting few empty cells by means of a formula described by Winer, Brown, & Michels, 1991). Mean RTs as a function of Naming Language, Stimulus Format and Number Magnitude are presented in Figure 5.

The main effect of Stimulus Format ($F(2, 38) = 2.66$, $MSE = 169791.1$, $p < .08$) tended towards significance. The effect of Naming Language was significant, $F(1, 19) = 171.53$, $MSE = 579952.4$, $p < .001$. Respective means for Dutch and 'Estonian' naming were 869 and 1339 ms. The interaction effect of Naming Language and Stimulus format was also significant, $F(2,$

38) = 541.74, *MSE* = 226321.0, $p < .001$.Tukey's HSD test showed that Forward translation (M = 1634 ms) was significantly slower than backward translation (M = 1406 ms), $p < .001$. Also, the effect of Number Magnitude ($F(14, 266) = 16.50$, *MSE* = 149204.6, $p < .001$) was significant, but all these effects were embedded in a three way interaction of Naming Language, Stimulus Format and Number magnitude, $F(28, 532) = 12.41$, *MSE* = 89479.4, $p < .001$. As can be seen in Figure 5 and similar to Experiments 1 and 2, the effect of Number Magnitude was only present in some of the Stimulus Format x Naming Language conditions. Again, these effects will be analyzed in detail in the following regression analyses.

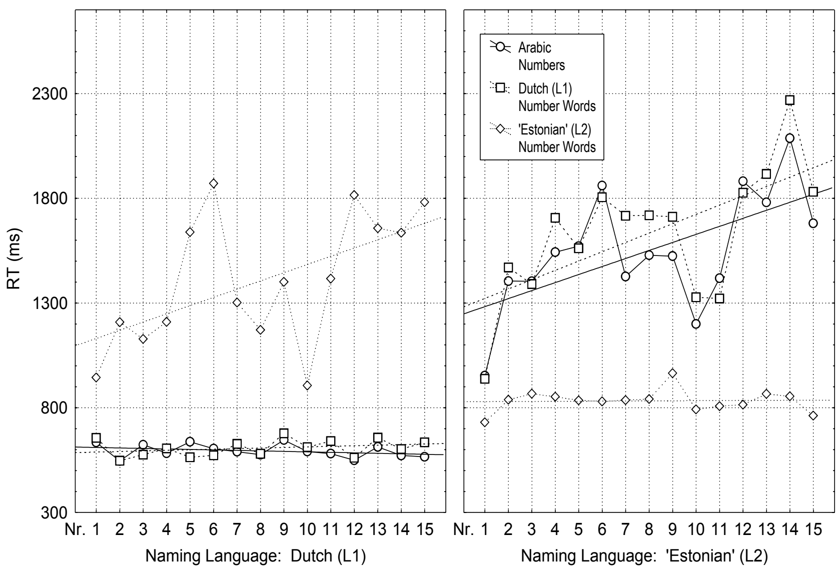


Figure 5. Mean naming RTs by naming language, stimulus format and number magnitude (Experiment 3 – 'Estonian' learning experiment). Straight lines represent best linear fit according to a least squares criterion.

Regression Analysis. The regression analyses were again performed by the procedure for repeated measures data described by Lorch and Myers (1990, Method 3). This approach has the additional advantage that missing data in the design (resulting from the use of an 85% learning criterion, see earlier) must not be estimated to carry out the analysis. For each participant, the multiple regression is calculated on the data that are available. Regression weights for the six Naming Language (L1 versus L2) x Stimulus Format (Arabic, L1, and L2) conditions are displayed in Table 3. Note that the effect of frequency could not be computed for the conditions with ‘Estonian’ stimulus words, as the participants had not encountered the words prior to the experiment, and each word was presented an equal number of times during the learning phase. Consequently, any effect of number magnitude in these conditions cannot be a word frequency effect. Similar to Experiment 1 and 2, the regression weights of Number Magnitude differed significantly from zero in both forward and backward translation conditions, respectively $t(19) = 2.427$, $p < .03$ and $t(19) = 2.157$, $p < .01$. These regression weights did not differ significantly from each other ($t(19) = 1.489$, $p > .15$), although the effect of number magnitude tended to be larger in the backward translation condition. Note that an effect of Frequency, supplementary to the effect of Number Magnitude, was found in the forward translation conditions, $t(19) = -3.407$, $p < .01$.

Naming Language	Stimulus Format		Intercept	Number Magnitude	Delayed Naming RT	Frequency
L1 Naming (Dutch)	Arabic Numbers	*Y* =	634	- 1.82 NM	- 0.29 D	+ 0.22 F
	L1 Number Words (Dutch)	*Y* =	748	+ 1.55 NM	- 0.57 D*	+ 0.15 F
	L2 Number Words ('Estonian')	*Y* =	963	+ 33.37 NM**	+ 0.26 D	
L2 Naming ('Estonian')	Arabic Numbers	*Y* =	2938	+ 16.03 NM(*)	+ 0.43 D	- 4.33 F**
	L1 Number Words (Dutch)	*Y* =	1219	+ 18.69 NM**	+ 5.69 D(*)	- 4.28 F**
	L2 Number Words ('Estonian')	*Y* =	142	- 0.30 NM	+ 2.10 D**	

Table 3. The regression equations for the six naming language x stimulus format conditions (Experiment 3) according to the procedure described by Lorch and Myers (1990) (* $p < .05$; ** $p < .01$; (*) p < .05 one-tailed test).

DISCUSSION

The results of Experiment 3 are quite surprising. Large semantic effects of number magnitude were found in both translation directions, even though the participants learned the 'Estonian' number words only a few minutes prior to the translation task. The obtained pattern of regression weights was very similar to that of Experiment 1 and 2 (with slower overall mean RTs). A comparison of Figure 5 with Figures 2 and 3 clearly illustrates this. Just as in Experiment 1, there was also a (somewhat weaker) effect of number magnitude for L2 naming of Arabic digits. The only remarkable difference with the previous experiments concerns the finding that forward translation was significantly slower than backward translation, while the reverse pattern was observed in Experiment 1 (and to a lesser extent in Experiment 2). Hence, this prediction of the hierarchical model holds at very low levels of L2 proficiency, while Experiment 1 shows that the opposite might be true for more proficient (though unbalanced) bilinguals. Note that we also found an effect of frequency in the forward translation condition, supplementary to the semantic number magnitude effect, suggesting a larger degree of lexical activation than in the previous experiments with more proficient bilinguals. However, this effect was not present in the backward translation condition; nor did not nullify the observed semantic number magnitude effects.

Because 'Estonian' L2 proficiency was extremely low in this experiment (participants saw each word only a few times), this strongly suggests that learned (number) word forms are mapped onto existing abstract (magnitude related) semantic information very early in the L2 acquisition process. Moreover, this semantic information is activated to a certain degree when translating the presented word forms. This is not a line with a strong developmental hypothesis which states that newly learned words are initially only mapped onto the lexical representation of their translation equivalents. We will go into further details about the theoretical implications of these results in the *General Discussion*.

Although our findings are surprising, they are not without analogues in the literature. In a study on the development of automatic processing, Tzelgov, Yehene, Kotler, and Alon (2000) taught participants a sequence of nonsense symbols, which represented magnitudes from 1 to 9 (although this was never told to the participants). Participants learned about the sequence of the symbols by indicating for pairs of symbols which one was the larger. After a limited amount of training, participants showed the classical effects associated with the number line. In particular, they showed the distance effect (i.e., they were faster to indicate the larger symbol when the distance between the magnitudes was large than when it was small), and they showed the physical size congruity effect (i.e., it was easier to select the physically larger symbol when this symbol represented a large magnitude than when it represented a small magnitude). Similarly, Logan and Klapp (1991) reported that participants could fluently verify equations of the form A + 2 = C (in which the digit indicated how many letters one had to go down the alphabet) after they had memorized 6 facts in a single session by means of rote rehearsal.

Finally, the fact that we obtained magnitude effects in number translation with stimulus words that were strongly controlled for their structure and the amount of prior exposure, is further evidence that the magnitude effect originates from the activation of semantic information, and is not a confound of the frequency with which different number words have been translated in the past, or any type of inherent structure that may be present in real-language number words.

EXPERIMENT 4

Now that we have shown that number translation is semantically mediated in both directions for a very wide range of L2 proficiency levels, one remaining question is to what extent the number magnitude effect is dependent on the restricted semantic context created by the experimental procedure itself. In the previous experiments, the same – limited – set of stimuli was presented several times in order to obtain reliable RTs in all conditions. In addition, all stimuli came from the same semantic category. This repeated presentation may have caused strong activation of the mental number line (see earlier), which may have increased the impact of the semantic system on the translation process.

There are reasons to doubt this possibility. For instance, if the number line had been excessively activated, this should also have caused a magnitude effect in the naming of number words, or indeed in the L1 naming of Arabic digits. The semantic route can be used for correct performance in these conditions as well (e.g. Fias, 2001; Reynvoet et al., 2002). Also, an analysis of the initial trials of Experiment 1 showed that the number magnitude effect was already present from the first few trials on[5].

Nevertheless, as the hierarchical model has mainly been developed to explain out-of-context translation performance in experimental settings, we wanted to investigate empirically whether our findings would still hold in a less restricted semantic context. Therefore, we designed a translation experiment in which each of the twelve number words was only presented

[5] We ran an analysis of only the first two (out of ten) trials of the first (out of two) naming block of Experiment 1. Hence, we only used the first 10% of all trials for each subject in our analysis (estimating the resulting few empty cells by means of a formula described by Winer et al., 1991). Despite the small number of datapoints included in this illustrative analysis, the regression weights of number magnitude were significant for both forward (B = 6.94, $t(21)$ = 5.135, p < .001) and backward translation (B = 5.08, $t(21)$ = 7.940, p < .001). Thus, the magnitude effect in number word translation was already present during the first trials of the experiment, at which point such long-term priming by semantic context may be argued to have not yet taken effect. The used blocked stimulus presentation sequence of Experiment 2 could not result in sufficient datapoints to perform a similar analysis, because the participants completed all 10 trials of a particular Naming Language x Stimulus Format condition before proceeding to the next block. As it seemed unjustifiable to use more than one Stimulus Format condition per subject in such an analysis, it would have been based on only two subjects per condition.

once. Moreover, these number words were scattered among 230 other middle to high frequency filler words which had to be translated as well. To keep the amount of number presentations as low as possible, and because the previous experiments showed consistent results for these conditions, we did not include within-language naming and Arabic digit naming. Consequently, the present study consisted of two subsequent blocks of forward and backward translation with respectively Dutch (L1) and French (L2) stimulus words. Thus, as in Experiment 2, the present experiment contained blocked stimulus language presentation.

METHOD

Participants. Twenty-nine first-year university students participated for course requirements. They all belonged to the group of bilinguals labelled before as 'unbalanced'.

Materials. All materials were identical to the previous experiments. The stimuli for each language consisted of the number words ranging from one to twelve and 230 other middle to high frequency filler words.

Design. The experiment had a 2 (Direction of Translation: Forward versus Backward) x 12 (Number Magnitude) full factorial design. All variables were manipulated within-subjects.

Procedure. All participants completed two blocks, each consisting of 242 trials. Hence, each (number word) stimulus was only presented once. The twelve number word trials were randomly scattered among the 230 filler trials. Naming language was Dutch (L1) in one block (French L2 stimuli), and French (L2) in the other (Dutch L1 stimuli). Order of blocks was determined at random. The procedure for a trial was identical to that used in Experiments 1, 2 and 3.

RESULTS

Variance Analysis. The proportion of invalid number word trials due to naming errors or faulty time registration was 17.0%. Because the nature of the design allowed only one observation per condition, trials were excluded from all analyses on-line as soon as there was the slightest doubt concerning accurate functioning of the voice key, in order to get mean RTs that were as reliable as possible with such a low number of observations. These trials were excluded from all analyses. An ANOVA was performed with Direction of Translation and Number Magnitude as repeated measures factors. The dependent variable was the mean RT across correct trials (estimating the resulting few empty cells by means of a formula described by Winer et al., 1991). Mean RTs as a function of Direction of Translation and Number Magnitude are presented in Figure 6.

The effect of Direction of Translation was nearly significant, $F(1, 28) = 3.84$, $MSE = 144547.4$, $p < .06$. Respective means for forward and backward translation were 1053 and 996 ms. As in the previous experiments, there was also a significant effect of Number Magnitude, $F(11, 308) = 10.32$, $MSE = 43349.1$, $p < .001$. This number magnitude effect just failed to interact with Direction of Translation, $F(11, 308) = 1.67$, $MSE = 42817.5$, $p < .08$. Again, these effects will be analyzed in detail in the following regression analyses.

Regression Analysis. The regression analyses were again performed by the procedure for repeated measures data described by Lorch and Myers (1990, Method 3). This approach has the advantage that the empty cells in the design, resulting from false trials combined with the use of a single observation per condition, must not be estimated to carry out the analysis. Regression weights for the two Direction of Translation conditions are displayed in Table 4. Similar to the previous experiments, the regression weights of Number Magnitude differed significantly from zero in both forward and backward translation conditions, respectively $t(28) = 2.097$, $p < .05$ and $t(28) = 2.403$, $p < .03$. The effect of Number Magnitude tended to be larger for forward translation, but this difference was not significant, $t(28) = 1.146$, $p > .25$. No lexical effects of word frequency were found.

		Intercept	Number Magnitude	Delayed Naming RT	Frequency
Backward Translation	*Y* =	1,721	+ 9.91 NM*	- 1.53 D	- 0.55 F
Forward Translation	*Y* =	2938	+ 24.37 NM*	- 0.16 D	+ 0.23 F

Table 4. The regression equations for the two direction of translation conditions (Experiment 4) according to the procedure described by Lorch and Myers (1990) (* $p < .05$; ** $p < .01$).

DISCUSSION

Again, we found an effect of number magnitude for both directions of translation. Consequently, it can be concluded that the translation of number words is semantically mediated also when these words are presented in a diversified semantic context. The fact that we obtained the number magnitude translation effect with one observation per participant per condition only, is a further indication that the effect is quite robust. The effect of number magnitude tended to be larger for forward translation (as predicted by the RHM), but this difference failed to reach significance. No effect of word frequency was found, suggesting a limited degree of lexical activation. Finally, it was interesting to note that, just as in Experiment 3 (very low L2 proficiency) but in contrast with Experiment 1 (and to a lesser extent with Experiment 2), backward translation was faster ($p < .06$, two-tailed) than forward translation. Hence, both L2 proficiency and semantic context were partly responsible for the opposite speed difference (faster forward translation) in Experiment 1. Note that these factors however, did not substantially affect the degree of semantic activation during translation.

GENERAL DISCUSSION

The RHM of Kroll and Stewart (1994; Kroll & de Groot, 1997), depicted in Figure 1, postulates that forward translation is more likely to be conceptually mediated than backward translation, because links between the lexicon and the semantic system are stronger for L1 than for L2. Backward translation will only be semantically mediated for high levels of L2 proficiency. As discussed in the introduction, these assumptions have been supported by a number studies (e.g. Kroll & Stewart, 1994; Talamas et al., 1999; Cheung & Chen, 1998; Sholl et al., 1995; for a review, see Kroll & de Groot, 1997; Kroll et al., 2002).

However, the results of our number word translation experiments were not completely in line with some of the model's predictions. In four experiments, we obtained clear semantic effects of number magnitude, not only when number words were translated forward from L1 to L2, but also when they were translated backward from L2 to L1. Thus, for Dutch-French bilinguals, it took less time to forward translate *twee* (L1) into *deux* (L2) [two] than *acht* (L1) into *huit* (L2) [eight], but also to backward translate *deux* into *twee* than *huit* into *acht*. There was no statistically reliable difference in the number magnitude effect between backward and forward translation. The observed difference slightly tended towards the expected direction for two of the four experiments (i.e., a smaller magnitude effect in backward translation than in forward translation for Experiments 1 and 4), but in the opposite direction for the other two experiments.

In addition, the effect of number magnitude in the translation conditions did not interact with L2 proficiency. It was equally strong in unbalanced and balanced bilinguals in Experiment 1, despite the reliable difference in mean RTs between both groups. Also, we replicated the number magnitude translation effect in participants with extremely low L2 proficiency with number words which had been acquired only a few minutes before the translation experiment (Experiment 3). This suggests that number word forms are mapped onto existing semantic information (the mental number line in this particular case) very early in the L2 acquisition process. Moreover, the mappings are strong enough to exert an influence on both backward and forward translation. These findings are not in line with a strong developmental hypothesis which states that L2 word forms only have connections with the abstract concepts which they represent at high levels of L2 proficiency. On the other hand, our observation of rapid semantic connections between new symbols and number magnitudes is in line with the literature of numerical cognition (see earlier Logan & Klapp, 1991; Tzelgov, Yehene, Kotler, & Alon, 2000).

We also found conflicting evidence concerning the relative speed of the two translation directions: forward translation was faster than backward translation in Experiment 1 and 2 (although the effect only reached significance in Experiment 1), while the opposite pattern was observed in Experiments 3 and 4, even when we used the same stimuli and equally proficient bilinguals (Experiment 4). The first finding is additional, indirect evidence that backward translation does not always rely on a faster lexical route. The second finding shows that both L2 proficiency and semantic context, which were manipulated in respectively Experiment 3 and 4, are determining factors for the relative speed of forward and backward translation. Such a finding is hard to explain within the current theoretical framework of the RHM, which claims that the relative speed of the translation routes only depends on L2 proficiency. It also shows that a faster backward translation process may not always be considered unambiguous evidence that the translation process was of a purely lexical nature.

There are a few other studies that failed to confirm the predictions of the RHM (Kroll & Stewart, 1994). Using a bilingual Stroop task, La Heij et al. (1996, Experiment 1) found that congruent colour words (for which the ink colour corresponded to the word), were translated faster than incongruent colour words. This was true for both directions of translation. In further experiments (La Heij et al., 1996, Experiments 3, 4 and 5), they found a facilitation effect (which was larger on translation trials than on naming trials) of distractor pictures depicting an object (e.g. a table) belonging to the same semantic category as the target word to be translated (e.g. *chair*), compared to unrelated pictures. This was also true for both directions of translation. Similar effects of context words and pictures in backward translation were recently reported by Bloem and La Heij (2003). But, as Kroll and de Groot (Kroll & de Groot, 1997, pp. 183-184) argued, context availability induced by the accompanying pictures may have provided the semantic support for L2 to access the meaning system, while the RHM was designed to account for out-of-context translation performance. It can be argued that this criticism also applies to our first three experiments. Although no explicit semantic cues (such as the distractor stimuli of La Heij et al. and Bloem and La Heij) were used in our experiments, repeated presentation of the same – limited – stimulus set may have induced sufficient semantic activation for L2 to access the semantic system (similar to results reported by Sholl et al., 1995). However, the increased semantic activation account is a much less plausible explanation for Experiment 4, where the number words were presented only once, amidst a whole range of other words from different semantic categories. Also, this account rests uneasily with the observation that the magnitude effect was present in translation conditions only, and not in direct word naming or in the L1 naming of Arabic digits, although the semantically mediated route also plays a role in these tasks (e.g. Fias, 2001; Reynvoet et al., 2002).

Another recent study suggesting early L2 lexico-semantic links, even after a few hours of artificially L2 learning, is that of Altarriba and Mathis (1997). After training a group of monolinguals on a set of English – Spanish word pairs, they found more errors on both lexically and semantically related false translations than on unrelated words in a translation recognition task (Altarriba & Mathis, 1997, Experiment 1). In a more recent study by von Pein and Altarriba (2003), similar findings were reported for English participants learning non-iconic American Sign Language gestures. In line with our findings in the 'Estonian' L2 learning study (Experiment 3), these findings suggest that links between L2 and the conceptual system can be established quite early. However, Altarriba and Mathis also reported more form related interference for nonfluent than for fluent bilinguals, supporting the developmental hypothesis of the RHM.

In further experiments (Altarriba & Mathis, 1997, Experiment 2), they reported a bilingual Stroop effect similar to that found by La Heij et al. (1996) using the same L2 training procedure as in their first experiment. Talamas et al. (1999) attributed these findings to the fact that the involved semantic representations may have been primed by the semantic procedures used during the training phase. Again, this alternative explanation by Talamas et al. would seem to be less applicable to our third experiment, given the fact that no semantic memorization strategy was used during the word learning phase. On the contrary, the associative word learning procedure that was used, was more likely to elicit L2 word learning by means of lexical connections.

Finally, evidence against translation asymmetries has been reported by de Groot and colleagues (e.g. de Groot, Dannenburg, & van Hell, 1994; Van Hell & de Groot, 1998b; de Groot & Poot, 1997; de Groot, 1992; de Groot & Comijs, 1995). Their translation studies showed that semantic word variables (e.g. imageability or context availability) affected response times both in backward and forward translation. Although there were some (although not always convincing) indications of asymmetries in the translation processes in some of the studies, this was not the case in the studies of de Groot and Comijs (1995) and de Groot and Poot (1997), where semantic variables affected forward and backward translation to an equal degree in three groups of bilinguals, differing in their level of L2 proficiency.

To conclude this discussion, we would like to point to some theoretical implications of the present findings. In line with the above studies, our results are clearly not compatible with a strong asymmetric model of bilingual memory, nor with a strong developmental lexical hypothesis.

Although the RHM has done a good job in explaining a wide range of findings and in producing new research hypotheses, the incompatible findings described above (as well as this study) indicate that some updating may be warranted. In our view, two amendments could entail a considerably useful extension of the model.

First, the model, like many models of that time, is implicitly based on the horserace metaphor. There are two routes, and the fastest wins (i.e., completely determines the output). So, translation either follows the lexical route (in backward translation) or the semantic route (in forward translation). There is no influence from the slower route. Increasingly, however, horse-race models are being replaced by connectionist-type models (the most famous example being Coltheart's dual route model of visual word naming; see Coltheart, 1978 vs. Coltheart, Curtis, Atkins and Haller, 1993). Following this approach, the central question should no longer be whether the output comes from one *or* the other route, but *how much* each of the routes contributes to the build-up of the overall output activation. In this view, one route is not faster than the other (the processing cycle is the same throughout the model); it only may have stronger connection weights and, therefore, influence the activation of the output units to a larger degree. If we apply this line of reasoning to the RHM, depicted in Figure 1, this would mean that the connection weights are stronger from L2 lexical units to L1 lexical units than the other way around. Similarly, the connection weights between L1 units and semantic units would be stronger than those between L2 units and semantic units. So, even though both the direct lexical-lexical and the indirect lexical-semantic-lexical route change the activation level of the units at each processing cycle, their relative contributions can differ (and maybe appear even nonexistent if the respective weights are very small) as a function of the translation direction.

A second proposal to improve the current theoretical framework of the RHM concerns the fact that the asymmetry only depends on the proficiency level of the bilingual. No distinction is made between different types of words.

Because of this, the model has to predict the same semantic involvement for the translation of all types of words, including number words, abstract words, and even syntactic function words. For this reason too, our findings with number words are critical for the model as a whole, even though there is independent evidence that the linking between new symbols and meanings is particularly fast for numerical stimuli (Logan & Klapp, 1991; Tzelgov et al., 2000). It would seem to us that the general framework of the RHM can easily be adapted to include influences at the word level as well as at the subject level (e.g. proficiency), certainly if a connectionist-type of model is used. For a start, the connection weights between the lexical units and the semantic units would depend on the consistency of the mappings between the words and the meanings. For each language, they would be bigger (and grow much faster in the acquisition phase) for words that always have the same meaning, independent of the context (e.g., *two*) than for words that have different meanings as a function of the context (e.g., *great*). Second, it does not seem unreasonable to assume that some words have a richer semantic representation than others, which could be implemented by the number of semantic features to which the word units are connected. Finally, the impact of the semantically mediated route on translation times would also depend on the degree of semantic overlap between two translation equivalents. Note that some of these ideas are already partly present in the distributed feature model of de Groot and colleagues (e.g. de Groot, 1992; de Groot, 1993; de Groot et al., 1994; Van Hell & de Groot, 1998a; Van Hell & de Groot, 1998b). According to this model, the overlap in meaning, indexed by the number of shared semantic features, depends on word concreteness[6] (with concrete translation equivalents sharing more features than abstract words). These shared features become active during translation, and

[6] Recent research (see Tokowicz & Kroll, 2003) has shown that certain word concreteness effects may alternatively be explained by competition effects resulting from cross-language ambiguity (i.e. the number of translation equivalents in each language for a given concept, which correlates with concreteness). This explanation of these effects is also quite compatible with the model depicted in Figure 7.

facilitate the translation process, resulting in faster translation of concrete words (e.g. de Groot, 1992; de Groot et al., 1994; Van Hell & de Groot, 1998b). This line of reasoning might also explain the strong semantic effects obtained in this study, as the meaning of number words is virtually completely overlapping across languages. At this point, it is important to note that the findings of Duyck, Szmalec, Kemps and Vandierendonck (2003), suggest that the early lexico-semantic mapping observed in Experiment 3, is probably not restricted to words from which the meaning is so clearly defined and overlapping as is the case for number words. Using a selective interference paradigm, they showed that new word forms are mapped onto available existing semantic (visual) information during associative word – new word learning (e.g. *auto* – *plornam* [car – legal Dutch nonword]), provided that such a visual representation is available. This shows that visual information – if possible - is coded early during the word acquisition process, which is compatible with the work of de Groot and colleagues.

Figure 7 gives the broad outline of how such a model could look like[7]. Note however, that the model described above is still a hypothetical description, which of course needs to be implemented before all the intricacies become clear. In the model, the semantic overlap is different for certain types of words (just as in the distributed feature model), with the overall weights of the lexico-semantic connections being stronger for L1 (just as in the RHM). This results in the activation of a smaller amount of semantic nodes, feeding activation into the L1 lexicon during backward translation of certain L2 words (but not for words with a very large semantic overlap between languages, e.g. *eight*).

---

[7] Notice that the model in Figure 7 contains one integrated lexicon. We refer to Dijkstra and Van Heuven (2002) for a recent review of evidence supporting this assumption.

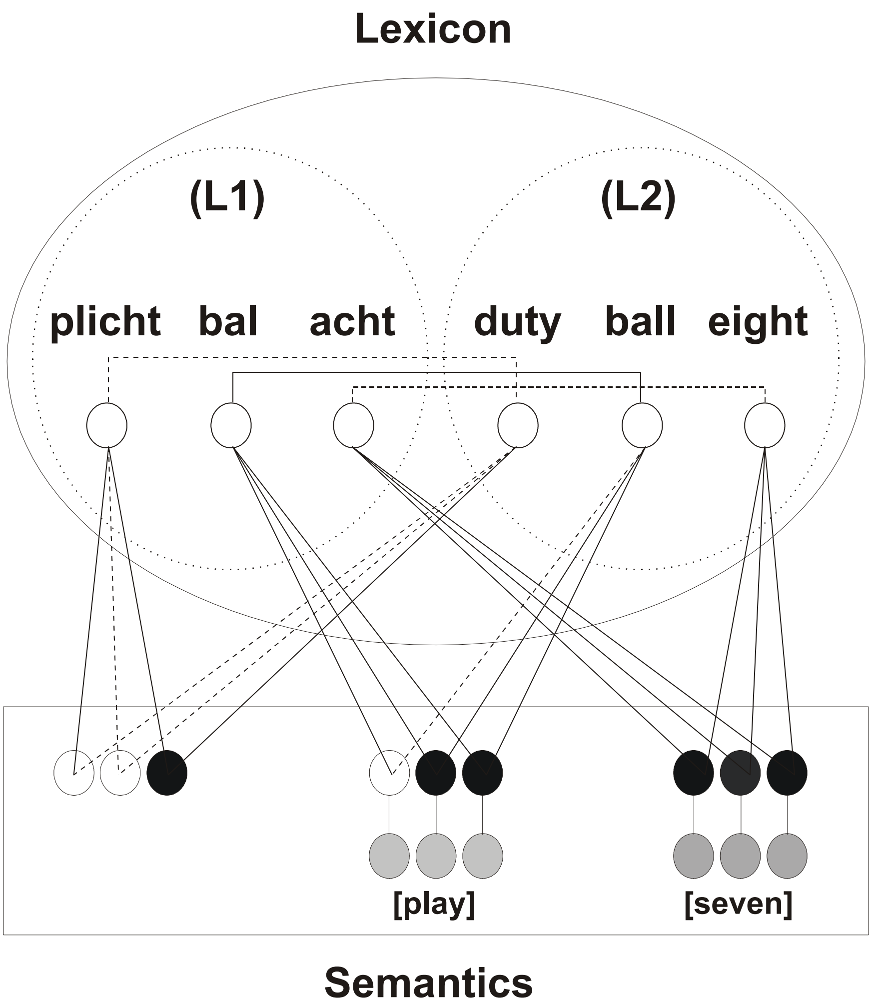


Figure 7. The RHM model of bilingual memory with varying semantic overlap and differently weighted lexico-semantic and intralexical connections (adapted from Kroll & de Groot, 1997). Solid lines represent stronger links than dotted lines. Depicted words and semantic representations are illustrative examples for Dutch-English bilinguals.

Notice that the model of Figure 7 also provides a straightforward interpretation of the semantic activation account, previously used by Kroll and de Groot (1997) to explain the data of the Stroop experiments (La Heij et al., 1996) for example (see earlier). In a connectionist-type model, the change of activation caused in a unit by a particular connection is not only a function of the weight of that connection, but also of the activation level of the unit where the connection originates from. So, enhancing the activation level of the semantic units (e.g., by the presentation of a picture) would already suffice to increase the impact of the semantic route in a model like the one depicted in Figure 7.

Finally, it might be argued that connection weights differ as a function of word characteristics not only between the lexical and the semantic level, but possibly also between the L1 and the L2 lexicon. It could be that these connections are stronger for words with a large form overlap (e.g., Figure 7: *ball* - *bal* for an English-Dutch bilingual) than for words with a small form overlap (e.g., Figure 7: *duty* - *plicht* for an English-Dutch bilingual). This could provide an explanation for the fact that words with a large form overlap (so-called cognates) are easier to translate and show less evidence for semantic mediation in the translation process, than words with no form overlap (e.g. de Groot, 1992; see also Sánchez-Casas et al., 1992). However, in the model described above, such strong lexical links do not necessarily exclude all semantic influences. This is compatible with Kroll and Stewart (1994), who demonstrated that cognate translation may be affected by the semantic organization of list context. Note that the lexico-lexical connections, depicted in Figure 7, between two word nodes are bidirectional (just as the lexico-semantic connections in the RHM and the intralexical connections in the BIA+ model, see Dijkstra & Van Heuven, 2002). This does not mean that the impact of these lexical connections will always be equally large for both directions of translation. On the contrary, the asymmetry (which is also present in the RHM) follows from the lexico-semantic weights which are weaker for certain L2 words (e.g. *duty*, *ball*) than for their L1 counterparts. This may cause smaller incoming semantic activation in the L1 word node during backward translation, resulting in a relatively larger impact of the intralexical activation (even with a bidirectional weight). However, at present we do not exclude the possibility of unidirectional connections with differing weights (as is the case in the

RHM at the lexical level). Future modeling will have to show the necessity of this assumption.

In summary, we found a semantic effect of number magnitude when number words were translated from Dutch (L1) to French (L2) and vice versa. Number words representing smaller quantities (e.g. *twee*, *deux* [two]) were translated faster than number words representing larger quantities (e.g. *acht*, *huit* [eight]). The effect was replicated using different procedures, semantic contexts, and different levels of L2 proficiency (including a very low level). These results strongly suggest that, at least for certain types of words, the mappings between L2 words and their meaning are more important than the lexico-lexical mappings between the L2 words and their L1 equivalents, already from the first stages of L2 acquisition on. On the basis of these findings, we have concluded that translation should not be viewed as an ‘all or none’ semantic or lexical process, but rather as the simultaneous build-up of activation from both the lexical and the semantic route. Furthermore, we have suggested that the contribution of each route not only depends on the translation direction and on the L2 proficiency, but also on the characteristics of the words involved in the translation. As such, we proposed a model in which the overall architecture of Kroll and Stewart's RHM is preserved, but in which we reviewed the way in which the different components interact, by dropping the implicit horserace assumption.

REFERENCES

Altarriba, J., & Mathis, K. M. (1997). Conceptual and Lexical Development in Second Language Acquisition. *Journal of Memory and Language, 36*(4), 550-568.

Altenberg, E. P., & Cairns, H. S. (1983). The Effects of Phonotactic Constraints on Lexical Processing in Bilingual and Monolingual Subjects. *Journal of Verbal Learning and Verbal Behavior, 22*(2), 174-188.

Bloem, I., & La Heij, W. (2003). Semantic Facilitation and Semantic Interference in Word Translation: Implications for Models of Lexical Access in Language Production. *Journal of Memory and Language, 48*(3), 468-488.

Brysbaert, M. (1995). Arabic Number Reading - on the Nature of the Numerical Scale and the Origin of Phonological Recoding. *Journal of Experimental Psychology: General, 124*(4), 434-452.

Brysbaert, M. (in press). Number Recognition in Different Formats. In J. I. D. Campbell (Ed.), *Handbook of Mathematical Cognition.* Psychology Press.

Brysbaert, M., Van Dyck, G., & Van de Poel, M. (1999). Visual Word Recognition in Bilinguals: Evidence From Masked Phonological Priming. *Journal of Experimental Psychology: Human Perception and Performance, 25*(1), 137-148.

Buckley, P. B., & Gillman, C. B. (1974). Comparisons of Digits and Dot Patterns. *Journal of Experimental Psychology, 103*(6), 1131-1136.

Campbell, J. I. D. (1994). Architectures for Numerical Cognition. *Cognition, 53*, 1-44.

Cheung, H., & Chen, H. C. (1998). Lexical and Conceptual Processing in Chinese-English Bilinguals: Further Evidence for Asymmetry. *Memory & Cognition, 26*(5), 1002-1013.

Cipolotti, L., & Butterworth, B. (1995). Toward a Multiroute Model of Number Processing - Impaired Number Transcoding With Preserved Calculation Skills. *Journal of Experimental Psychology: General, 124*(4), 375-390.

Coltheart, M. (1978). Lexical Access in Simple Reading Tasks. In Underwood, G. (Ed.), *Strategies of Information-Processing.* (pp. London: Academic Press.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of Reading Aloud - Dual-Route and Parallel-Distributed-Processing Approaches. *Psychological Review, 100*(4), 589-608.

de Groot, A. M. B. (1992). Determinants of Word Translation. *Journal of Experimental Psychology: Learning, Memory and Cognition, 18*(5), 1001-1018.

de Groot, A. M. B. (1993). Word-type Effects in Bilingual Processing Tasks. Support for a Mixed-Representational System. In Schreuder, R. and Weltens, B. (Eds.), *The Bilingual Lexicon: Vol. 6. Studies in Bilingualism.* (pp. 27-57). Amsterdam, Netherlands: John Benjamins Publishing Company.

de Groot, A. M. B., & Comijs, H. (1995). Translation Recognition and Translation Production - Comparing a New and an Old Tool in the Study of Bilingualism. *Language Learning, 45*(3), 467-509.

de Groot, A. M. B., Dannenburg, L., & van Hell, J. G. (1994). Forward and Backward Word Translation by Bilinguals. *Journal of Memory and Language, 33*(5), 600-629.

de Groot, A. M. B., & Poot, R. (1997). Word Translation at Three Levels of Proficiency in a Second Language: the Ubiquitous Involvement of Conceptual Memory. *Language Learning, 47*(2), 215-264.

Dehaene, S. (1992). Varieties of Numerical Abilities. *Cognition, 44*(1-2), 1-42.

Dehaene, S., & Akhavein, R. (1995). Attention, Automaticity, and Levels of Representation in Number Processing. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*(2), 314-326.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The Mental Representation of Parity and Number Magnitude. *Journal of Experimental Psychology: General, 122*(3), 371-396.

Dehaene, S., Dupoux, E., & Mehler, J. (1990). Is Numerical Comparison Digital - Analogical and Symbolic Effects in 2-Digit Number Comparison. *Journal of Experimental Psychology: Human Perception and Performance, 16*(3), 626-641.

Dehaene, S., & Mehler, J. (1992). Cross-Linguistic Regularities in the Frequency of Number Words. *Cognition, 43*(1), 1-29.

Dijkstra, A., & Van Heuven, W. (2002). The Architecture of the Bilingual Word Recognition System: From Identification to Decision. *Bilingualism: Language and Cognition, 5*(3), 175-197.

Dijkstra, T., Grainger, J., & Van Heuven, W. J. B. (1999). Recognition of Cognates and Interlingual Homographs: the Neglected Role of Phonology. *Journal of Memory and Language, 41*(4), 496-518.

Dijkstra, T., Timmermans, M., & Schriefers, H. (2000). On Being Blinded by Your Other Language: Effects of Task Demands on Interlingual Homograph Recognition. *Journal of Memory and Language, 42*(4), 445-464.

Duyck, W., & Brysbaert, M. (2002). What Number Translation Studies Can Teach Us About the Lexico-Semantic Organisation in Bilinguals. *Psychologica Belgica, 42*(3), 151-175.

Duyck, W., Szmalec, A., Kemps, E., & Vandierendonck, A. (2003). Verbal Working Memory Is Involved in Associative Word Learning Unless Visual Codes Are Available. *Journal of Memory and Language, 48*(3), 527-541.

Ferrand, L. (1999). Why Naming Takes Longer Than Reading? The Special Case of Arabic Numbers. *Acta Psychologica, 100*(3), 253-266.

Fias, W. (2001). Two Routes for the Processing of Verbal Numbers: Evidence from the SNARC Effect. *Psychological Research,*

Fias, W., Lammertyn, J., Reynvoet, B., Dupont, P., & Orban, G. A. (2003). Parietal Representation of Symbolic and Nonsymbolic Magnitude. *Journal of Cognitive Neuroscience, 15*(1), 47-56.

Fias, W., Reynvoet, B., & Brysbaert, M. (2001). Are Arabic Numerals Processed as Pictures in a Stroop Interference Task? *Psychological Research, 65*(4), 250-259.

Foltz, G. S., Poltrock, S. E., & Potts, G. R. (1984). Mental Comparison of Size and Magnitude - Size Congruity Effects. *Journal of Experimental Psychology: Learning, Memory and Cognition, 10*(3), 442-453.

Gielen, I., Brysbaert, M., & Dhondt, A. (1991). The Syllable-Length Effect in Number Processing Is Task- Dependent. *Perception & Psychophysics, 50*(5), 449-458.

Ito, Y., & Hatta, T. H. (2003). Semantic Processing of Arabic, Kanji, and Kana Numbers: Evidence From Interference in Physical and Numerical Size Judgments. *Memory & Cognition, 31*(3), 360-368.

Jescheniak, J. D., & Levelt, W. J. M. (1994). Word-Frequency Effects in Speech Production - Retrieval of Syntactic Information and of Phonological Form. *Journal of Experimental Psychology: Learning, Memory and Cognition, 20*(4), 824-843.

Koechlin, E., Naccache, L., Block, E., & Dehaene, S. (1999). Primed Numbers: Exploring the Modularity of Numerical Representations With Masked and Unmasked Semantic Priming. *Journal of Experimental Psychology: Human Perception and Performance , 25*(6), 1882-1905.

Kroll, J. F. (1993). Assessing Conceptual Representations for Words in a Second Language. In Schreuder, R. and Weltens, B. (Eds.), *The Bilingual Lexicon: Vol. 6. Studies in Bilingualism.* (pp. 53-82). Amsterdam, Netherlands: John Benjamins Publishing Company.

Kroll, J. F., & de Groot, A. M. B. (1997). Lexical And Conceptual Memory in the Bilingual: Mapping Form to Meaning in Two Languages. In de Groot, A. M. B. and Kroll, J. F. (Eds.), *Tutorials In Bilingualism: Psycholinguistic Perspectives.* (pp. 201-224). Mahwah, NJ: Lawrence Erlbaum Publishers.

Kroll, J. F., Michael, E., Tokowicz, N., & Dufour, R. (2002). The Development of Lexical Fluency in a Second Language. *Second Language Research, 18*, 137-171.

Kroll, J. F., & Stewart, E. (1994). Category Interference in Translation and Picture Naming - Evidence for Asymmetric Connections Between Bilingual Memory Representations. *Journal of Memory and Language, 33*(2), 149-174.

La Heij, W., Hooglander, A., Kerling, R., & Vandervelden, E. (1996). Nonverbal Context Effects in Forward and Backward Word Translation: Evidence for Concept Mediation. *Journal of Memory and Language, 35*(5), 648-665.

Logan, G. D., & Klapp, S. T. (1991). Automatizing Alphabet Arithmetic .1. Is Extended Practice Necessary to Produce Automaticity. *Journal of Experimental Psychology: Learning, Memory and Cognition, 17*(2), 179-195.

Lorch, R. F., & Myers, J. L. (1990). Regression-Analyses of Repeated Measures Data in Cognitive Research. *Journal of Experimental Psychology: Learning, Memory and Cognition, 16*(1), 149-157.

McCloskey, M. (1992). Cognitive Mechanisms in Numerical Processing: evidence from Acquired Dyscalculia. *Cognition, 44*, 107-157.

Meuter, R. F. I., & Allport, A. (1999). Bilingual Language Switching in Naming: Asymmetrical Costs of Language Selection. *Journal of Memory and Language, 40*(1), 25-40.

Moyer, R. S., & Landauer, T. K. (1967). Time Required for Judgments of Numerical Inequality. *Nature, 215*, 1519-1520.

Naccache, L., & Dehaene, S. (2001). The Priming Method: Imaging Unconscious Repetition Priming Reveals an Abstract Representation of Number in the Parietal Lobes. *Cerebral Cortex, 11*(10), 966-974.

Nas, G. (1983). Visual Word Recognition in Bilinguals - Evidence for a Cooperation Between Visual and Sound Based Codes During Access to a Common Lexical Store. *Journal of Verbal Learning and Verbal Behavior, 22*(5), 526-534.

Pesenti, M., Thioux, M., Seron, X., & De Volder, A. (2000). Neuroanatomical Substrates of Arabic Number Processing, Numerical Comparison, and Simple Addition: a Pet Study. *Journal of Cognitive Neuroscience, 12*(3), 461-479.

Pinel, P., Dehaene, S., Riviere, D., & Lebihan, D. (2001). Modulation of Parietal Activation by Semantic Distance in a Number Comparison Task. *Neuroimage, 14*(5), 1013-1026.

Reynvoet, B., & Brysbaert, M. (1999). Single-Digit and Two-Digit Arabic Numerals Address the Same Semantic Number Line. *Cognition, 72*(2), 191-201.

Reynvoet, B., Brysbaert, M., & Fias, W. (2002). Semantic Priming in Number Naming. *Quarterly Journal of Experimental Psychology, 55A*(4), 1127-1139.

Sánchez-Casas, R. M., Davis, C. W., & Garcia-Albea, J. E. (1992). Bilingual Lexical Processing - Exploring the Cognate Non- Cognate Distinction. *European Journal of Cognitive Psychology, 4*(4), 293-310.

Sholl, A., Sankaranarayanan, A., & Kroll, J. F. (1995). Transfer Between Picture Naming and Translation - a Test of Asymmetries in Bilingual Memory. *Psychological Science, 6*(1), 45-49.

Smith, M. C. (1997). How Do Bilinguals Access Lexical Information? In de Groot, A. M. B. and Kroll, J. F. (Eds.), *Tutorials In Bilingualism: Psycholinguistic Perspectives.* (pp. 145-167). Mahwah, NJ: Lawrence Erlbaum Publishers.

Talamas, A., Kroll, J. F., & Dufour R. (1999). From Form to Meaning: Stages in the Acquisition of Second-Language Vocabulary. *Bilingualism: Language and Cognition, 2*(1), 45-58.

Thomas, M. S. C., & Allport, A. (2000). Language Switching Costs in Bilingual Visual Word Recognition. *Journal of Memory and Language, 43*(1), 44-66.

Tokowicz N., & Kroll J. F. (2003). *Cross-language ambiguity and concreteness effects: The role of meaning and form ambiguity in cross-language processing*. Manuscript submitted for publication.

Tzelgov, J., Yehene, V., Kotler, L., & Alon, A. (2000). Automatic Comparisons of Artificial Digits Never Compared: Learning Linear Ordering Relations. *Journal of Experimental Psychology: Learning, Memory and Cognition, 26*(1), 103-120.

Van Hell, J. G., & de Groot, A. M. B. (1998a). Conceptual Representation in Bilingual Memory: Effects of Concreteness and Cognate Status in Word Association. *Bilingualism: Language and Cognition, 1*, 193-211.

Van Hell, J. G., & de Groot, A. M. B. (1998b). Disentangling Context Availability and Concreteness in Lexical Decision and Word Translation. *Quarterly Journal of Experimental Psychology, 51A*, 41-63.

Van Wijnendaele, I., & Brysbaert, M. (2002). Visual Word Recognition in Bilinguals: Phonological Priming from the Second to the First Language. *Journal of Experimental Psychology: Human Perception and Performance, 3*, 619-627.

von Pein M., & Altarriba J. (2003). *The Development of Linguistic Knowledge in Naïve Learners of American Sign Language*. Manuscript submitted for publication.

Von Studnitz, R. E., & Green, D. W. (1997). Lexical Decision and Language Switching. *International Journal of Bilingualism, 1*, 3-24.

Winer, B.J., Brown, D.R., & Michels, K.M. (1991). *Statistical Principles in Experimental Design*. (3rd ed.). Boston, Massachusetts: McGraw-Hill.

Chapter 3
Verbal Working Memory Is Involved in Associative Word Learning Unless Visual Codes Are Available

Journal of Memory and Language, 2003, 48, 527-541[1, 2]

Baddeley, Gathercole and Papagno (1998) proposed a model of associative word learning in which the phonological loop, as defined in Baddeley's working memory model, is primarily a language learning device, rather than a mechanism for the memorization of familiar words. Using a dual-task paradigm, Papagno, Valentine and Baddeley (1991) found that articulatory suppression, loading verbal working memory, had an effect on the memorizing of word-nonword pairs, but not on the memorizing of word-concrete word pairs. The present work explored the potential for visual codes in unfamiliar word learning. In a first experiment, we replicated the results of Papagno et al. for both nonwords and highly imageable nouns. In addition, we found that articulatory suppression disrupted the memorizing of word-abstract word pairs, suggesting that phonological involvement may be triggered by the absence of visual representations for the abstract words. Experiment 2 showed that an artificially induced association between a nonword and a non-nameable visual image was sufficient to compensate for diminished verbal working memory resources due to articulatory suppression. In a third experiment, we demonstrated that our results generalize to other types of abstract words (i.e. function words), auditory stimulus presentation, and to word learning in children.

---

[1] This paper was co-authored by Arnaud Szmalec, Eva Kemps and André Vandierendonck

[2] The authors are indebted to Brian McElree, Lucy Henry, Wouter De Baene, and two anonymous reviewers for critical comments on an earlier draft of this article.

INTRODUCTION

The present paper addresses the importance of visual codes for the acquisition of new words by means of associative learning in adults and children. As learning new vocabulary is crucial to intellectual development (Sternberg, 1987), the identification of the cognitive processes involved in acquiring new words is of major importance. This can lead to important theoretical and practical insights concerning language acquisition both in healthy adults and children, and in patients with language impairments.

Psycholinguistic research on language acquisition has focused on the association between concepts and words (e.g. Markman, 1994) or on how the syntax of a language is adopted (e.g. Gleitman, 1993). Another line of research has dealt with working memory (WM) involvement during the first stages of word acquisition (for an extensive review, see Baddeley, Gathercole, & Papagno, 1998). This work is characterized by an emphasis on the acquisition of the phonological representation of new words. The WM model developed by Baddeley and Hitch (Baddeley, 1986; Baddeley & Hitch, 1974) has been shown to constitute an appropriate theoretical framework to investigate the role of phonological codes in word acquisition. The model comprises three components: a central executive (CE) and two subsidiary slave systems, the phonological loop (PL) and the visuo-spatial sketch pad (VSSP). The CE serves as an attentional control mechanism, and is responsible for coordinating the operations of the two slave systems. The PL is responsible for the short-term storage and processing of verbal material, such as spoken words. It can also provide verbal encodings of visually presented material such as written words and nameable pictures. Rapid decay of the phonological representations in the store can be offset by a strategic rehearsal process. The phonological store is operational from the age of three (Gathercole & Adams, 1993), while the rehearsal process is fully developed only after the age of seven (Gathercole & Hitch, 1993; see also Henry & Millar, 1991, 1993; Kemps, De Rammelaere, & Desmet, 2000). The VSSP is involved in the temporary retention and manipulation of visuo-spatial material, such as spatial patterns and locations.

This WM model has been successfully incorporated within neuropsychological and developmental areas of research (Baddeley, 1997). Moreover, there is substantial support for the neural substrates of the PL (e.g. Baddeley, Papagno, & Vallar, 1988; Grasby, Frith, Friston, Bench, Frackowiak, & Dolan, 1993; Paulesu, Frith, & Frackowiak, 1993). Also, the WM model provides an attractive theoretical framework when using dual-task methodology: the involvement of a particular WM system in a given task can be investigated by comparing performance under single-task and dual-task conditions. If primary task performance is affected by simultaneous execution of a dual-task that loads only one of the components, it can be assumed that the WM component involved is necessary for the execution of the primary task. It should also be noted that a revised WM model was recently proposed to include a third slave system, the episodic buffer (Baddeley, 2000). This component is believed to function as a temporary interface between the slave systems and long term memory.

A substantial body of evidence has been accumulated for a model of language acquisition which proposes that the verbal component of Baddeley's WM model (the PL) is primarily involved in the storage of unfamiliar sound patterns of words until more stable (long-term) representations can be established. Hence, it is not longer viewed as a mechanism for the memorization of familiar words. Therefore, the PL has been designated as "primarily a language learning device" (Baddeley et al., 1998). The next sections summarize several relevant studies for this hypothesis, categorized by participants (children vs. adults), design (correlational vs. experimental) and language (native vs. foreign) (for a more extensive overview, see Baddeley et al., 1998).

Developmental research has demonstrated the importance of a verbal WM system such as the PL for word acquisition in children. Positive correlations have been observed between measures of verbal WM capacity and native vocabulary knowledge in children of various ages, particularly when nonword repetition scores were used to measure verbal WM capacity instead of the more widely used digit span[3] (Bowey, 2001; Gathercole & Adams, 1993, 1994; Gathercole, Hitch, Service, & Martin, 1997; Gathercole, Service, Hitch, Adams, & Martin, 1999; Gathercole, Willis, Emslie, & Baddeley, 1992; Michas & Henry, 1994). Service (1992) for example showed that nonword repetition scores were a significant predictor of the performance of Finnish children learning English two years later. Cheung (1996) and Masoura and Gathercole (1999) replicated this finding with respectively Chinese and Greek children learning English. As for experimental studies, Gathercole and Baddeley (1990) showed that children with low nonword repetition scores performed more poorly on a native word learning task than children with higher verbal WM capacity. Native word learning was experimentally conceptualized as learning the association between unfamiliar names (e.g. *Pimas*) and toy animals. There was no difference between groups with respect to learning the association between the same toys and familiar names (e.g. *Thomas*). Similar findings were reported by Gathercole and al. (1997) and Michas and Henry (1994), who also showed that experimental word learning performance is positively correlated with phonological memory skills. These findings suggest that the link between verbal WM capacity and native vocabulary acquisition remains after controlling for language exposure.

---

[3] This is probably due to the fact that nonwords, unlike digits, do not have any semantic or lexical representation, which can affect memory span scores (Bourassa & Besner, 1994; Hulme, Maughan, & Brown, 1991; Poirier & Saintaubin, 1995; Wetherick, 1975). However, even nonword repetition scores may not be a pure measure of phonological storage capacity, as they are also influenced by language-specific probability of used phonotactic segments (Gathercole, Frankish, Pickering, & Peaker, 1999)

With regard to research on word learning by adults, Papagno and Vallar (1992) showed that phonological similarity and item length had an effect on the associative memorizing of word-nonword pairs, but not on the memorization of word-word pairs. Both these effects of phonological similarity and word length are attributed to the operation of verbal WM: phonologically similar and longer items require more verbal WM resources than phonologically dissimilar and shorter items. Hence, Papagno and Vallar's findings suggest involvement of verbal WM in associative new word learning. Later, Papagno and Vallar (1995) found that polyglots have greater digit spans and higher nonword repetition scores than control participants. Polyglots were also better at memorizing word-nonword pairs, but not at memorizing word-word pairs, even though they did not perform better on tests assessing general intelligence or visuo-spatial abilities.

Direct experimental evidence for verbal WM involvement in language learning by adults comes from a study by Papagno, Valentine and Baddeley (1991). Following classical dual-task logic, they reasoned that it should be possible to show verbal WM involvement in associative word learning by demonstrating an interference effect of a simultaneously performed task which loads the PL. In contrast, such a task should not interfere with performance on an associative word memorization task if both words are known, because this association can be learned using other codes (i.e. semantic, visual, …) than phonological ones. Using Italian participants, Papagno et al. found that articulatory suppression (repeated uttering of the sound "bla"), a secondary task known to load the PL, interfered with the learning of Italian-Russian pairs (e.g. *libro-cniga*), but not of Italian word pairs (e.g. *lupo-carta*). This was found for both auditory and visual stimulus presentation. A replication with English participants learning English-Finnish (e.g. *cowboy-pila*) and English-English (e.g. *roof-hunter*) word pairs yielded similar results. However, the researchers failed to replicate the effect for English-Russian word pairs (Experiments 3 and 4). They claimed that this might be due to the fact that the participants succeeded in learning the Russian words under articulatory suppression by making use of lexical or semantic associations. For example, the word pair *throat-garlo* may have been learned by lexically associating *garlo* with the English word *gargle*, which in turn can be associated semantically with *throat*.

Recent neuropsychological studies support the assumption that some (e.g. semantic) variables originating from long-term memory may influence performance in verbal WM tasks. Hanten and Martin (2001) showed that BS, a patient with a developmental phonological short-term memory deficit, was able to perform well in a wide range of learning and memory tasks if he could make use of lexical or semantic information. However, his performance dropped significantly if this was not possible, such as for learning lists of words of low frequency and low imageability. Similarly, Martin and Saffran (1999) found that the ability of aphasic patients with lexical and short-term memory deficits to learn supraspan word lists (i.e. lists of which the length exceeds the patients' working memory capacity) was influenced by word imageability, word frequency and the linguistic relationship between the words of a list. This is similar to the imageability effect on word repetition performance which is typically observed in patients suffering from deep dysphasia, a rare language impairment associated with a phonological short-term memory deficit (Majerus, Lekeu, Van Der Linden, & Salmon, 2001). Also, Bird, Franklin and Howard (2002) showed that the discrepancies between nouns and function words in comprehension and production performance of aphasic patients disappeared when imageability was controlled.

In summary, there is a substantial body of evidence in support of the involvement of a verbal WM system such as the PL in learning new native or foreign vocabulary until more stable long-term representations are formed (Baddeley et al., 1998). However, people use existing (e.g. semantic, lexical, …) long-term language knowledge to mediate verbal learning whenever possible (Papagno et al., 1991, Experiment 3 and 4; Hanten & Martin, 2001; Martin & Saffran, 1999).

Most of the studies mentioned above are of a correlational nature and therefore provide only indirect evidence for the involvement of the PL in vocabulary learning. The possibility that a third causal factor accounts for the common variance in the two associated constructs cannot be ruled out. For example, it is possible that an enriched linguistic environment (e.g. better education, exposure to books, …) results in a larger vocabulary and a greater working memory capacity. Hence, the observed relation between word learning and working memory capacity may not be a causal one. Of course, this criticism does not apply to the handful of experimental word learning studies (Gathercole & Baddeley, 1990; Gathercole et al., 1997; Michas & Henry, 1994).

The study of Papagno et al. (1991) provided more direct evidence for verbal WM involvement in foreign word acquisition. However, their results are subject to a methodological constraint. Just as in the two studies of Papagno and Vallar (1992, 1995), word imageability was not taken into account when selecting stimuli: almost all target words were highly imageable[4]. Therefore, not only did word-word pairs differ from word-nonword pairs with regard to the novelty of the second word in the pair, but also with regard to the availability of a strong link between the second word and a visual representation. By definition, this was not the case for the nonwords. This confound might explain the absence of an articulatory suppression effect on the learning of the word-word pairs: participants might have used a visual memorization strategy (e.g. imagining a picture of a wolf with a card in its mouth for the *lupo-carta* pair) to learn the word-word associations, whereas only a verbal strategy was available for the word-nonword pairs, due to the

[4] For all but one of their word stimuli (i.e. *attic*), an imageability rating could be found in the MRC psycholinguistic database (Coltheart, 1981; Fearnley, 1997; Wilson, 1988): mean imageability values were 585 (Experiment 1 and 2; *SD* = 37), 572 (Experiment 3 and 4; *SD* = 58) and 589 (Experiment 6; *SD* = 42), measured on a imageability scale ranging from 100 to 700. Mean imageability was only moderate for the stimuli of Experiment 7 (i.e. 364, *SD* = 44)

absence of a visual representation for the nonwords. An associative imageable word pair memorization paradigm, in which two words are encoded by means of an image combining the visual representations of both words, is typically used in studies that elicit visual memorization as a method to investigate VSSP functioning (Andrade, Kemps, Werniers, May, & Szmalec, 2002; e.g. Logie, 1986; Quinn & McConnell, 1996). This line of reasoning is in agreement with the neuropsychological studies of Bird et al. (2002) Hanten and Martin (2001), Majerus et al. (2001) and Martin and Saffran (1999) mentioned earlier, who reported effects of imageability on verbal WM performance of patients with various phonological short-term memory deficits. Similar beneficial effects of word imageability on verbal short-term memory performance of people without such a deficit have been reported by Paivio and Smythe (1971) and Walker and Hulme (1999).

In conclusion, it is possible that the findings of Papagno and colleagues can be attributed to the fact that the participants used a different memorization strategy (i.e. imagery) for the word-word pairs. If this is true, then verbal working memory is not involved in word learning because the words are new, but because they do not yet have a strong association with any visual representation.

This confounding factor, however, does not entirely minimize the importance of verbal working memory in novel word learning. The effect of articulatory suppression on the learning of nonwords suggests that verbal working memory is indeed involved in word learning, but it is possible that the locus of its involvement is limited to the learning of novel phonological representations; the learning of the word associations themselves can rely on other (e.g. visual) WM resources. This hypothesis will be investigated in the following experiments, and further discussed in more detail in the *General Discussion* section. Hence, we believe that, although the learning of phonological codes is important in vocabulary acquisition, it should not be restricted to this aspect, because semantic and visual representations are probably equally important.

EXPERIMENT 1

To investigate the involvement of visual codes in the learning of word pairs, the present study used an associative word learning experiment with word pairs of which the second word was low in imageability. If, in accordance with Papagno et al. (1991), verbal WM is involved only in the associative memorization of word-new word pairs, articulatory suppression should not affect performance on these pairs. However, it is our view that such an effect would occur due to the fact that abstract (low imageable) words are not strongly associated with any visual representation. It should also be noted that Papagno et al. always compared performance under articulatory suppression with performance under concurrent matrix tapping (a secondary task loading the VSSP). We believe it is more useful to compare performance under articulatory suppression with a single task condition, in order to get a purer indication of the effect of diminished verbal WM resources. No such control condition was included in the study of Papagno et al..

METHOD

Participants. Forty-eight first-year students enrolled at the Faculty of Psychology and Educational Sciences, Ghent University, participated for course requirements and credit. Their native language was Dutch.

Design. The experiment was a 3 (target word: concrete, abstract, nonword) x 2 (suppression: control, articulatory suppression) x 5 (trial: one to five) design. Target word was included as a between-subjects factor, while suppression and trial were manipulated within subjects. The number of correctly recalled target words (from zero to eight) was the dependent variable.

Materials. All words were chosen from Van Loon-Vervoorn (1985), who obtained imageability ratings for 4600 Dutch nouns on a seven point scale.

Two lists of word pairs were constructed for each of the three types of target words: one list for the control condition and another for the articulatory suppression condition. Lists were counterbalanced over the two conditions. Each list consisted of eight word pairs (see Appendix A). Each word pair consisted of a cue and a target word. The cue words were common Dutch nouns and were used to initiate the recall of the target words. The target words were concrete words, abstract words or nonwords which had to be remembered after presentation of the accompanying cue word. In all three target word conditions, the same cue words were used to ensure that differences between conditions were solely due to the target words. Both cue and target words consisted of two syllables. All cue words were highly imageable (M = 6.65, SD = 0.20). Cue words and target words could not be easily associated, either semantically (e.g. *roof-house*) or lexically (e.g. *roof-room*), so as to prevent problems as those encountered by Papagno et al. (1991, Experiments 3 and 4, see earlier).

All targets had moderate word frequency according to the CELEX counts (Baayen, Piepenbrock, & Van Rijn, 1993). Mean target word frequency was matched as closely as possible, to ensure that the concrete (high imageable) words were not more frequent than the abstract (low imageable) target words ($t < 1$).The mean log frequency per million of the cue words was 1.35 (SD = 0.70).

Word-concrete word pairs (high imageable target word). All target words were highly imageable (List One: M = 6.77, SD = 0.17; List Two: M = 6.68, SD = 0.06) nouns. The mean log frequency per million of the target words was 0.94 (SD = 0.67).

Word-abstract word pairs (low imageable target word). Cue words in this condition were the same as in the condition mentioned previously. All abstract target words were nouns with a very low imageability rating (List One: M = 1.80, SD = 0.20; List Two: M = 1.84, SD = 0.20). The mean log frequency per million of the target words was 1.00 (SD = 0.54).

Word-nonword pairs. Cue words in this condition were the same as in the other conditions. The nonwords were disyllabic strings of random vowels and consonants, chosen in such a way that they contained morphemes which are likely to occur in Dutch, but did not resemble existing Dutch words.

Procedure. Participants were randomly assigned to one of the three target word conditions (concrete, abstract or nonword), in which the two lists of eight word pairs were presented: one in the control condition and one in the articulatory suppression condition (in a counterbalanced order). Each participant was seated in front of a 15” screen, connected to an IBM compatible PC. The computer driven experiment started after extensive oral instructions. The procedure was as similar as possible to that of Papagno et al. (1991). Each trial consisted of a learning and a test phase. During the learning phase, the eight word pairs were presented centered on the screen in a random order. The cue word was presented above the target word. The pairs remained on the screen for two seconds, with a two second inter-trial-interval (ITI). Participants were asked to memorize the words, so that they would be able to recall the second word, after presentation of the first word. No indication was given concerning possible memorization strategies. During the test phase, all cue words were presented sequentially in a random order. Participants were required to type the appropriate word completely within a seven second interval. Then, the following cue word was presented. Each trial consisted of this learning and test phase. Each participant completed five of these trials in both the control and articulatory suppression conditions. In the latter, participants were asked to continuously utter the word ‘de’ (Dutch for ‘the’) during the learning phase. Suppression started four seconds before presentation of the first word pair and terminated four seconds after the last pair had been presented. Articulatory suppression was performed only during the encoding of the words, not during the test phase. The experiment lasted approximately 35 minutes.

RESULTS

The number of correctly recalled words was subjected to a 3 (target word: concrete, abstract, nonword) x 2 (suppression: control, articulatory suppression) x 5 (trial: one to five) ANOVA. Tests of analyses by participants and by items will be referred to as F_1 and F_2 respectively. A response was rated as 'correct' when it sounded like the correct word when it was pronounced according to Dutch grapheme-to-phoneme conversion rules. All means are displayed in Figure 1.

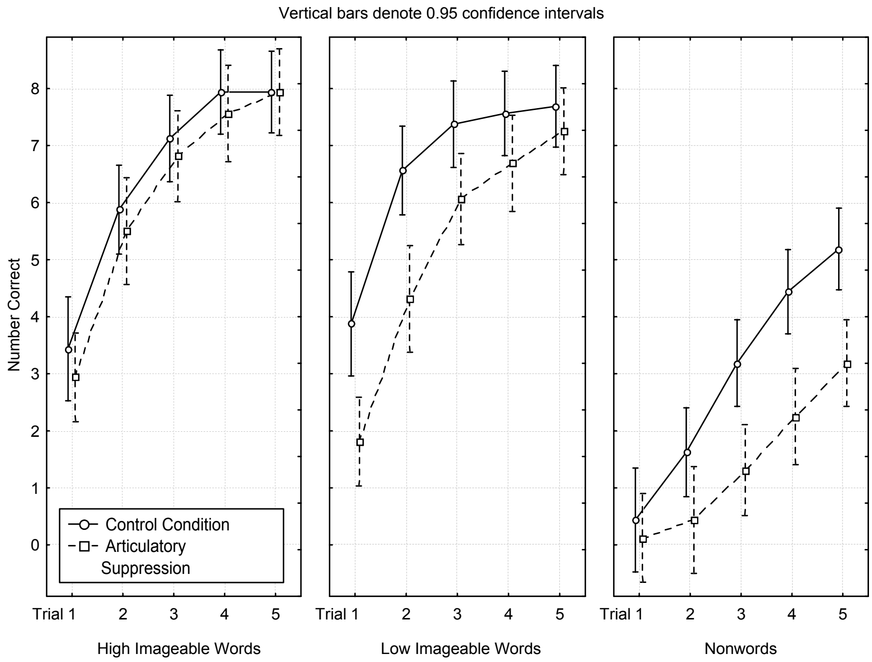


Figure 1. Mean number of correctly recalled target words by target word, suppression and trial (Experiment 1).

The main effect of target word was significant, $F_1(2, 45) = 67.94$, $MSE = 11.99$, $p < .001$, $F_2(2, 45) = 299.92$, $MSE = 2.93$, $p < .001$. Post-hoc comparisons using Tukey's HSD test showed that performance was significantly lower for word-nonword pairs than for word-concrete word and word-abstract word pairs (all p's < .001 for analyses by participants and by items). There was no significant difference between word-concrete word and word-abstract word pairs, $p_1 > .58$ and $p_2 > .16$. The effect of suppression was also significant, $F_1(1, 45) = 45.24$, $MSE = 3.04$, $p < .001$ and $F_2(1, 45) = 92.67$, $MSE = 1.86$, $p < .001$, just as the main effect of trial, $F_1(4, 180) = 175.33$, $MSE = 1.74$, $p < .001$ and $F_2(4, 180) = 348.71$, $MSE = 0.84$, $p < .001$. Tukey's HSD test showed significant differences ($p < .001$) between all trials except for trials four and five, which differed only in the analysis by items ($p_1 > .11$ and $p_2 < .02$). Hence, it seems that memorization performance began to level off somewhat after four trials.

As expected, the interaction between target words and suppression reached significance, $F_1(2, 45) = 5.72$, $MSE = 3.04$, $p < .01$, $F_2(2, 45) = 12.09$, $MSE = 1.86$, $p < .001$. More important, a planned comparison of the interaction involving only concrete and abstract words was also significant: the articulatory suppression effect was much stronger for word-abstract word pairs, $F_1(1, 45) = 7.60$, $MSE = 3.04$, $p < .01$, $F_2(1, 45) = 12.11$, $MSE = 1.86$, $p < .01$. Accordingly, planned comparisons showed a significant articulatory suppression effect for both abstract and nonwords, respectively $F_1(1, 45) = 25.32$, $MSE = 3.04$, $p < .001$, $F_2(1, 45) = 43.56$, $MSE = 1.86$, $p < .001$ and $F_1(1, 45) = 30.08$, $MSE = 3.04$, $p < .001$, $F_2(1, 45) = 70.47$, $MSE = 1.86$, $p < .001$. There was no articulatory suppression effect for concrete words, $F_1(1, 45) = 1.28$, $MSE = 3.04$, $p > .26$ and $F_2(1, 45) = 1.28$, $MSE = 3.04$, $p > .10$.

DISCUSSION

We succeeded in replicating the main findings of Papagno et al. (1991). Articulatory suppression disrupted associative word-nonword learning, suggesting verbal WM involvement in the acquisition of new words. Such an effect was not present when the task involved two highly imageable familiar words. However, articulatory suppression did also affect performance when the target was familiar, but had a low imageability rating. Hence, the conclusions of Papagno et al. regarding associative learning of familiar words should be restricted to highly imageable words. This supports the hypothesis that the absence of a visual code is the determining factor for verbal WM involvement in associative word learning, rather than the novelty of a word. It follows that the association between two words (not the respective phonological representations) may be learned by means of other than verbal (e.g. visual) WM resources.

EXPERIMENT 2

If the availability of a visual code is indeed the crucial factor counteracting the negative effects of articulatory suppression on the learning of word-concrete word pairs, imagery (in a visual working memory component such as the VSSP) may play a role in this kind of word learning. To test this hypothesis more directly, a second experiment was designed. We decided not to use visuo-spatial suppression as a secondary task to study the role of visual working memory in the learning of word-concrete word pairs for two reasons. First, any visuo-spatial suppression effect is likely to be circumvented through verbal memorization strategies. Such a strategy cannot be hindered by induction of articulatory suppression because it is undesirable to use two secondary tasks at the same time. Second, most active VSSP tasks are spatial rather than visual in nature (e.g. spatial tapping, Farmer, Berman, & Fletcher, 1986), while the passive secondary tasks are mainly visual (e.g. dynamic visual noise, Quinn & McConnell, 1996). No active, predominantly visual secondary task was found to be appropriate for this study. The method that we decided to use circumvents these problems. It seeks to remove the articulatory suppression effect on the learning of word-nonword pairs by means of inducing an association between the nonword and a non-nameable visual code. A thoroughly learned associated visual nonword representation may allow visuo-spatial WM resources to compensate for diminished verbal WM capacity imposed by the secondary verbal task. If this is the case, then the articulatory suppression effect can be expected to disappear. Then, it follows that imageability can be put forward more confidently as the crucial factor for verbal WM involvement in associative word learning.

METHOD

Participants. Sixteen first-year students enrolled at the Faculty of Psychology and Educational Sciences, University of Ghent, participated for course requirements and credit. They were all native Dutch speakers. None of them participated in Experiment 1.

Design. The experiment was a 2 (induction: control, visual code) x 2 (suppression: control, articulatory suppression) x 5 (trial: one to five) design. The factor induction was manipulated between subjects, while condition and trial were within-subjects factors.

Materials. The cue words and the nonwords were those used in Experiment 1. Sixteen non-nameable, monochrome line drawings were constructed and randomly assigned to the nonwords. Computer images drawn by hand were used to avoid clear geometrical figures (lines, triangles, …) which can easily be named. They are displayed in Appendix B.

Procedure. All participants were randomly assigned to one of the induction conditions. They received the two lists of word-nonword pairs: one to be memorized in the control condition and one under articulatory suppression (counterbalanced with order).

Each participant was seated in front of a 15” screen, connected to an IBM compatible PC. Instructions were presented on the screen. The experiment consisted of an association phase, a learning phase, and a test phase. During the association phase, participants in the visual induction condition had to learn the association between the nonwords and their corresponding visual codes. The line drawings were presented in a white square (169 cm²) on a black background. The corresponding nonword was presented above the white square in which the line drawing was presented. Participants in the control condition only saw these nonwords, with a white, empty square underneath. These stimuli (nonwords with or without visual codes) were all presented 20 times for a period of four seconds with a one second ITI. In the learning phase, word-nonword pairs were presented and memorized following the procedure of Experiment 1, both with and without articulatory suppression. In the test phase, memorization of the word pairs was tested as described in the procedure section of Experiment 1. The experiment lasted approximately 60 minutes.

RESULTS

The number of correctly recalled words was subjected to a 2 (induction: control, visual code) x 2 (suppression: control, articulatory suppression) x 5 (trial: one to five) ANOVA. Tests of analyses by participants and by items will be referred to as F_1 and F_2 respectively. A response was rated as 'correct' if it sounded like the correct word when it was pronounced according to Dutch grapheme-to-phoneme conversion rules. All means are displayed in Figure 2.

We observed significant main effects of suppression and trial, $F_1(1, 14) = 14.36$, $MSE = 5.96$, $p < .01$, $F_2(1, 15) = 27.53$, $MSE = 1.50$, p < .001 and $F_1(4, 56) = 66.90$, $MSE = 1.28$, $p < .001$; $F_2(4, 60) = 126.17$, $MSE = 0.37$, p < .001, respectively. Tukey's HSD test showed significant differences ($p < .05$) between all trials except between trials four and five in the analysis by participants ($p_1 > .23$ and $p_2 < .05$). Hence, it seems that memorization performance began to level off a bit after four trials. The effect of induction was only significant in the analysis by items, $F_1(1, 14) = 2.13$, $MSE = 17.36$, $p > .16$ and $F_2(1, 15) = 14.97$, $MSE = 1.30$, p < .01.

As expected, there was a significant suppression by induction interaction, $F_1(1, 12) = 4.77$, $MSE = 6.72$, $p < .05$ and $F_2(1, 15) = 10.60$, $MSE = 1.49$, p < .01. Planned comparisons revealed that the articulatory suppression effect

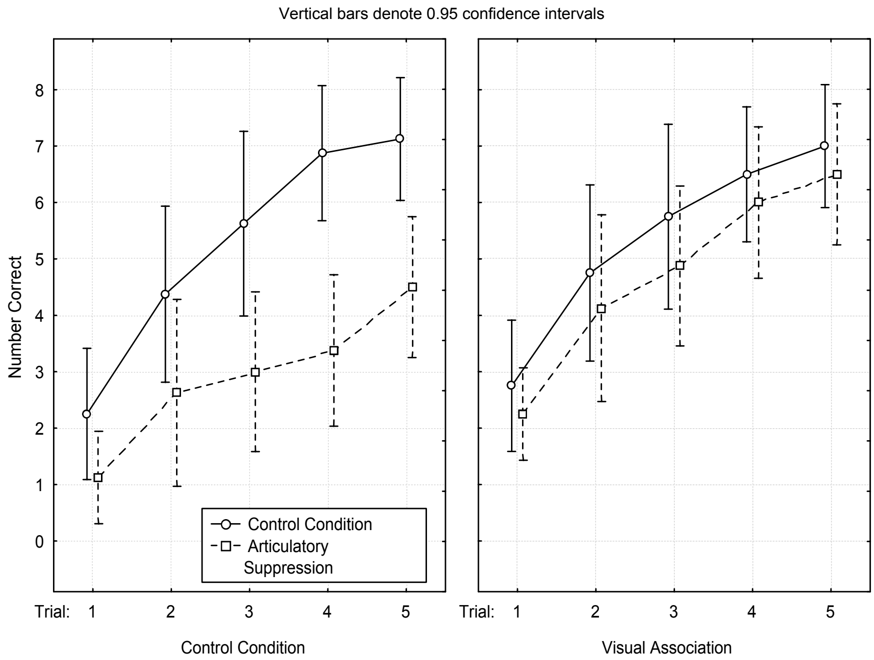


Figure 2. Mean number of correctly recalled target words by induction, suppression and trial (Experiment 2).

was not significant on any trial in the induction condition (all p_1's > .28; all p_2's > .18), nor across trials, $F_1(1, 14) = 1.21$, $MSE = 5.96$, $p > .29$ and $F_2(1, 15) = 3.03$, $MSE = 1.00$, $p > .10$. However, in the control condition, we observed a significant suppression effect on every trial (all p_1's < .02; all p_2's < .01), except for the first (probably due to a floor effect), $p_1 < .06$ and $p_2 < .05$. The effect of suppression was also significant across trials, $F_1(1, 14) = 18.15$, $MSE = 5.96$, $p < .001$, $F_2(1, 15) = 27.17$, $MSE = 1.99$, $p < .001$.

DISCUSSION

Experiment 2 confirms the hypothesis that the effect of articulatory suppression on associative word learning can be circumvented by artificially establishing an association between a nonword and a visual representation of a non-nameable line drawing. Articulatory suppression did not affect the learning of word-nonword pairs when participants had previously seen those nonwords twenty times together with their corresponding visual images. In the control condition without visual association, however, a verbal suppression effect was observed on every trial (except for a probable floor effect in the analysis by participants on the first trial). Hence, associative word learning only relies on verbal WM if a visual representation is not available. These results also suggest that the association between the nonwords and the line drawings was not learned via a verbal label (e.g. *zigzag*) assigned to the drawings, since this would probably have triggered an articulatory suppression effect.

It follows that the imageability, rather than the novelty of a word determines verbal WM involvement in learning associations between words. Therefore, it is plausible that the lack of verbal WM resources (due to articulatory suppression) can be compensated by using visual short-term memory strategies (such as imagery) when learning the association between pairs of words that have links with some visual representation. In agreement with Baddeley et al. (1998) and Papagno et al. (1991), the present experiments confirm that verbal WM is important when learning new native and foreign vocabulary. However, they also clarify that this phenomenon can be attributed to the absence of visual representations for new words, and that verbal WM may be necessary for learning phonological representations, but not for learning the word associations themselves.

EXPERIMENT 3

In this last experiment, we seek to investigate whether our findings regarding the importance of visual codes for the acquisition of words generalize to other a) age groups, b) types of words, and c) presentation modalities.

First, it is important to show that the effect of imageability on associative word learning is not only present in adults, because children learn more vocabulary than adults do. For example, it is estimated that pupils acquire around seven words per day (almost 3000 words per year) during the elementary through high school years (Nagy & Anderson, 1984; Nagy & Herman, 1987). Furthermore, because most studies on verbal working memory involvement in word learning by children are of a correlational nature (see earlier), it is useful to test the experimental word learning paradigm used in Experiment 1 in a younger age group. Because most Belgian (Dutch speaking) school children begin to learn English, French, and sometimes German, ancient Greek and Latin in the first year of high school, we decided to investigate the effect of imageability on word learning in a group of first year high school students (+/- 12 years old). Furthermore, a dual-task methodology such as the one used in Experiment 1, would be too demanding for younger children.

Second, because at least some 12 year olds may not know some of the abstract words used in Experiment 1 (e.g. *inteelt* [inbreeding]), we sought to generalize our previous findings to other types of low imageable words. We chose function words (e.g. *because, when, therefore,* ...) because these are learned at an early age and are by definition among the least imageable of all words. The fact that function words are also very frequent, and thus easier to remember, strengthens a potential effect of articulatory suppression on the associative learning of word-function word pairs.

Finally, we decided to use auditory stimulus presentation in the present experiment, to exclude the possibility that visual codes are only important in associative word learning when the words are presented visually.

METHOD

Participants. Forty-two pupils enrolled in the first year of the Klein Seminarie high school of Roeselare, Belgium, volunteered for this experiment. Their ages ranged from 11 years, 11 months to 13 years (*M* = 12.04; *SD* = 0.38). They were all native Dutch speakers.

Design. The experiment was a 3 (target word: concrete, abstract, nonword) x 2 (suppression: control, articulatory suppression) x 5 (trial: one to five) design. Target word was included as a between-subject factor while suppression and trial were manipulated within subjects. The number of correctly recalled target words (from zero to six) was the dependent variable.

Materials. Analogous to Experiment 1, two lists of six word pairs were constructed for each of the three target word conditions, by removing two items from the original stimuli of Experiment 1 (see Appendix A). This was done because a pilot study had indicated that learning eight word pairs was too difficult for children of this age. Again, the two lists were counterbalanced over the two suppression conditions.

The remaining cue words had a mean imageability of 6.66 on a seven point scale (*SD* = 0.21) according to the ratings reported by Van Loon-Vervoorn (1985). Their CELEX (Baayen et al., 1993) mean log frequency per million was 1.46 (*SD* = 0.71). The remaining concrete target words were all highly imageable (*M* = 6.70, *SD* = 0.14). As mentioned earlier, we chose function words as the abstract target words for this experiment, because a) it was likely that some of the abstract words of Experiment 1 were not well known by some children, b) function words are among the least imageable words and c) we sought to generalize our findings to other word types. Because no imageability ratings are available for Dutch function words, we considered the ratings for the English translation of those function words according to the MRC Psycholinguistic Database (Coltheart, 1981; Wilson, 1988; Fearnley, 1997). This assumption of cross-linguistic imageability similarity of translation equivalents is supported by the high (r = .95) correlation between the Dutch and the English MRC imageability ratings of the cue and target words of Experiment 1. The mean imageability of the function words for which a rating could be found was only 244.22 (*SD* = 25.20) on a scale from 100 to 700. The concrete target words were more imageable ($p < .001$), but less frequent ($p < .001$) than the abstract target words. The fact that function words are highly frequent words does not alleviate, but even strengthens a potential effect of articulatory suppression on the associative learning of word-function word pairs.

Word-concrete word pairs (high imageable target word). All target words were highly imageable (List One: *M* = 6.75, *SD* = 0.19; List Two: *M* = 6.65, *SD* = 0.04). Their mean log frequency per million was 0.93 (*SD* = 0.51).

Word-abstract word pairs (low imageable target word). The function words had very low imageability ratings (List One: *M* = 251.20, *SD* = 33.06; List Two: *M* = 235.50, *SD* = 7.33). Their mean log frequency per million was 2.26 (*SD* = 0.55).

*Word-nonword pairs***.** Just as in Experiment 1, the nonwords were disyllabic strings of random vowels and consonants, chosen in such a way that they contained morphemes which are likely to occur in Dutch, but did not resemble existing Dutch words.

Procedure. The procedure was identical to that of Experiment 1, but differed with regard to the presentation modality: words were not presented visually, but auditorily by means of headphones, using the timing parameters of Experiment 1. The subjects wrote down their responses in a notebook.

RESULTS

The number of correctly recalled words was subjected to a 3 (target word: concrete, abstract or nonword) x 2 (suppression: control and articulatory suppression) x 5 (trial: one to five) ANOVA. Tests of analyses by participants and by items will be referred to as F_1 and F_2 respectively. A response was rated as correct as soon as it sounded like the correct target word, according to Dutch grapheme-to-phoneme conversion rules. One participant was excluded from all analyses because he could not remember a single word, in any of the conditions. All means are displayed in Figure 3.

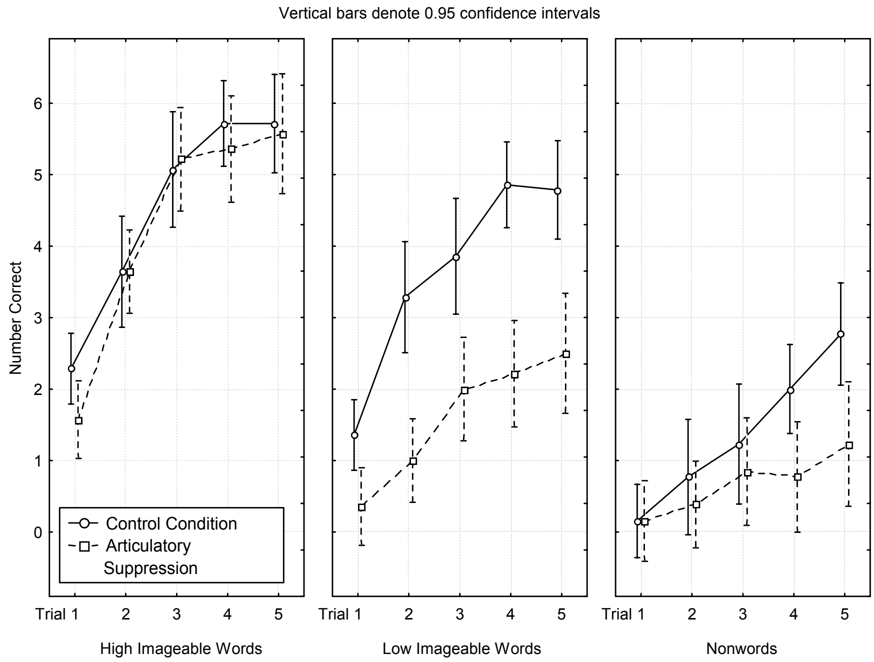


Figure 3. Mean number of correctly recalled target words by target word, suppression and trial (Experiment 3).

We observed a main effect of target word, $F_1(2, 38) = 50.36$, $MSE = 7.52$, $p < .001$, $F_2(2, 33) = 152.96$, $MSE = 3.13$, $p < .001$. Post-hoc comparisons using Tukey's HSD test indicated that performance was poorer for the abstract words than for the concrete words, and for the nonwords compared to the abstract words (all p's $< .001$ for analyses by participants and by items). The main effects of suppression, $F_1(1, 38) = 37.23$, $MSE = 2.63$, $p < .001$, $F_2(1, 33) = 23.94$, $MSE = 4.73$, $p < .001$, and trial, $F_1(4, 152) = 117.57$, $MSE = 0.90$, $p < .001$, $F_2(4, 132) = 232.78$, $MSE = 0.53$, $p < .001$ were also significant. Tukey's HSD test showed significant differences ($p < .001$) between the first three trials, but not between trials three and four, and four and five, respectively $p_1 > .19$, $p_2 < .05$ and $p_1 > .70$, $p_2 > .38$. Hence, it seems that memorization performance began to level off somewhat after three trials.

As expected, the interaction between target word and suppression was significant: $F_1(2, 38) = 11.46$, $MSE = 263$, $p < .001$, $F_2(2, 33) = 7.59$, $MSE = 4.73$, $p < .01$. More important, a planned comparison of the interaction involving only the concrete word and the abstract word conditions showed that the articulatory suppression effect was much weaker for the former than for the latter, $F_1(1, 38) = 21.52$, $MSE = 2.63$, $p < .001$, $F_2(1, 33) = 13.97$, $MSE = 4.73$, $p < .001$. Planned comparisons indicated that the suppression effect was significant for both the abstract words, $F_1(1, 38) = 53.90$, $MSE = 2.63$, $p < .001$, $F_2(1, 33) = 35.00$, $MSE = 4.73$, $p < .001$, and the nonwords: $F_1(1, 38) = 6.18$, $MSE = 2.63$, $p < .05$, $F_2(1, 33) = 3.72$, $MSE = 4.73$, $p < .065$. Articulatory suppression did not affect the memorization of concrete words, $F_1 < 1$, $F_2 < 1$. Finally, the effect of articulatory suppression was stronger for the abstract word condition than for the nonword condition, $F_1(1, 38) = 10.92$, $MSE = 2.63$, $p < .01$, $F_2(1, 33) = 7.94$, $MSE = 4.73$, $p < .01$, but this is probably due to a floor effect in the nonword condition.

DISCUSSION

All the main findings of Experiment 1 were replicated. As expected, articulatory suppression disrupted the associative learning of both word-nonword and word-function word pairs, but not of word-concrete word pairs. Therefore, the present experiment showed that our findings regarding the importance of visual codes for word learning can be generalized with respect to age (children and adults), word type (function words and nouns) and presentation modality (auditory and visual stimulus presentation). This is further evidence that the absence of a visual code is the determining factor for verbal WM involvement in associative word learning, rather than the novelty of the words.

GENERAL DISCUSSION

Following extensive evidence for verbal WM involvement in foreign and native vocabulary learning in both children and adults (e.g. Baddeley et al., 1998), we hypothesized that associative learning of (a) word-concrete word pairs would not be impaired by articulatory suppression, whereas memorization of (b) word-abstract word pairs and (c) word-nonword pairs would. Our data from Experiments 1 and 3 supported this hypothesis. These findings suggest that the conclusion of a number of studies showing that verbal WM is not involved in the associative word learning of familiar words (Baddeley, 1993; Baddeley et al., 1998; e.g. Papagno et al., 1991; Papagno & Vallar, 1992, 1995), only applies to familiar words which are highly imageable. The articulatory suppression effects for the abstract words in this study (nouns and function words) showed that verbal WM resources can be important for the associative learning of familiar words if the absence of visual representations for these words prevents the use of visual WM strategies such as imagery. This issue has been overlooked in previous studies.

Our hypotheses were further tested in Experiment 2, in which we showed that verbal WM involvement in the learning of word-nonword pairs may be minimized by associating a visual image with the nonword. Therefore, word imageability, rather than word novelty, appears to be the key factor that determines the degree of verbal WM involvement in associative word learning. It follows that verbal WM involvement in vocabulary acquisition is merely a consequence of the absence of visual codes for new words.

Baddeley et al. (1998) and Papagno et al. (1991) already mentioned the possibility that verbal WM involvement in associative word learning can be influenced by lexical or semantic factors after they failed to find an articulatory suppression effect on the learning of English-Russian word pairs (Papagno et al., 1991, Experiments 3 and 4, see earlier). Bourassa and Besner (1994) provided evidence that imageability, rather than other semantic or lexical long-term memory variables, is a key factor when investigating influences of long-term knowledge on verbal WM functioning. They found that serial ordered recall was better for content words than for function words. However, differences between the two word classes disappeared when the two stimulus sets were matched for word imageability. Accordingly, Walker and Hulme (1999) found that both backward and forward (written and spoken) serial recall was better for concrete words than for abstract words. These studies are in accordance with the beneficial effects of word imageability on verbal working memory performance observed in neurological patients with verbal short-term memory impairment (Bird et al., 2002; Hanten & Martin, 2001; Majerus et al., 2001; Martin & Saffran, 1999).

While our findings point to a methodological constraint of all previous experimental studies on verbal WM involvement in associative word learning, they do corroborate the importance of verbal WM in vocabulary acquisition. However, it is important to indicate precisely the locus of verbal WM involvement in that process. Freedman and Martin (2001; see also Martin, 1993; Martin, Shelton, & Yaffee, 1994) showed that there are dissociable phonological and semantic short-term memory components which are linked with corresponding representations in long-term memory. Like Freedman and Martin, we agree with Baddeley et al. (1998) that verbal WM is primarily a language learning device, but only if language learning is defined as the long-term learning of novel phonological forms. Hence, the PL (the phonological short-term memory component in Martin's terminology) is important in language acquisition, but only with respect to forming its corresponding (phonological) long-term representations. Although the learning of phonological codes is important in language learning, the concept of language learning should not be restricted to this aspect, because semantic and visual representations are probably equally important and sometimes acquired earlier than the corresponding phonological representations. A baby for example, has semantic and visual representations of its *mother* long before it acquires the phonological label for that concept. We therefore agree with Freedman and Martin (2001) that the impact of dissociable short-term memory components, such as the PL, on other semantic and visual long-term representations is limited.

This line of reasoning applies to the results of Experiment 2. Our findings do not rule out verbal WM involvement in the learning of nonwords in the visual induction condition. No doubt verbal rehearsal played a role during the association phase when both the phonological code of the nonword and its association with the visual code were learned (solid lines on the right hand side of Figure 4). During the learning phase, the participants learned the association between the cue words and the target words by keeping the two respective visual representations together in visual short-term memory (lower dotted lines), because articulatory suppression made it difficult to learn the association by keeping both phonological codes (e.g. [baik]-[pu:sti]) in verbal short-term memory (upper dotted lines). During the test phase, the phonological code of the cue word (e.g. [baik]) successively activated the visual representation of *bike,* the non-nameable visual image associated with *poosti,* and the phonological code of the target word (e.g. [pu:sti]).

In conclusion, it is important to make a distinction between the learning of word associations and the learning of phonological representations of new words. Verbal WM is crucial in language learning, but it is only important for the long-term learning of phonological representations. Our results clearly show that the locus of verbal WM involvement is not necessarily situated in the learning of the word associations themselves. The association can be learned by other means (e.g. a visual WM component such as the VSSP), while the phonological representations cannot.

This study has some practical implications for vocabulary learning in children, for adults learning a foreign language and for word learning in neuropsychological patients. The results demonstrate facilitation effects due to the availability of visual codes when learning new words. Such an effect

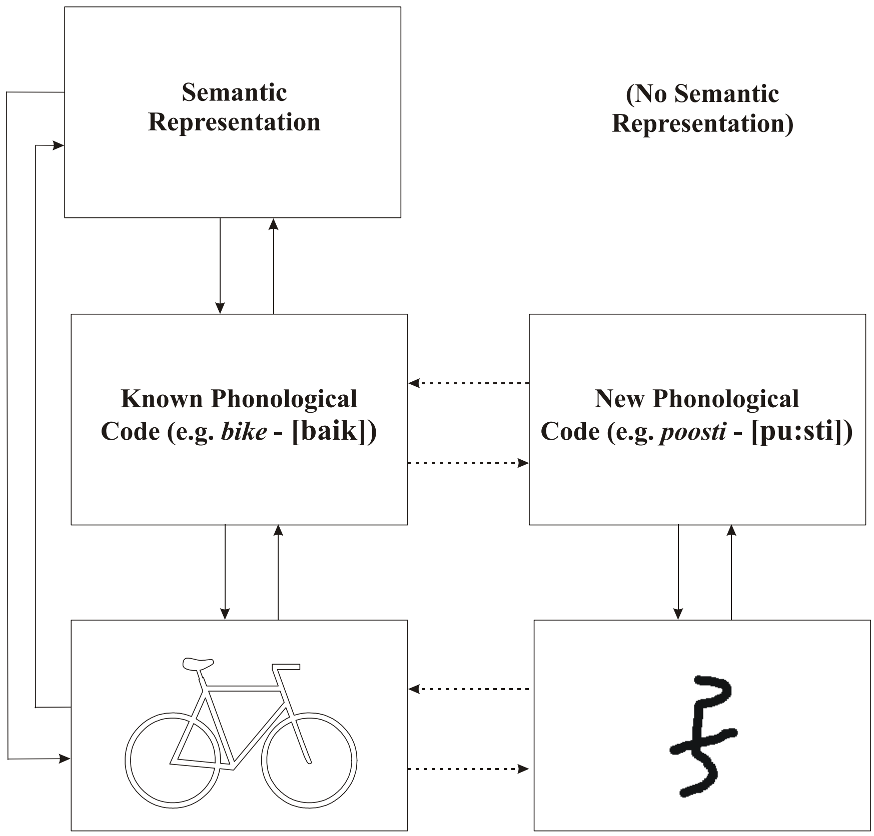


Figure 4. An extended model of Freedman and Martin's (2001) view on learning word-new word associations including visual codes.

may be especially important when semantic or phonological representations have not yet fully developed, as is the case for very young children learning their first words. For example, it may be useful to point to items when teaching a young child a new concrete word. Similarly, foreign language learning in adults may benefit from the use of pictorial material or visual imagery mnemonics. This hypothesis is supported by research on bilingualism, which has shown that links between second language words and semantic information are established quite early during the learning process (Altarriba & Mathis, 1997; Duyck & Brysbaert, 2004), in contrast with assumptions of earlier models of bilingual language organization (e.g. Kroll & Stewart, 1994). As for neuropsychological patients, providing visual information may be sufficient to compensate for verbal short-term memory deficits. Hanten and Martin (2001) showed that BS, a patient suffering from a substantial phonological short-term memory impairment, but who nonetheless obtained a PhD in biology, performed very well in a variety of learning and memory tasks, provided he could use lexical and semantic information. If this was not possible, such as for learning lists of words of low frequency and low imageability, his performance dropped significantly. Similar beneficial effects of imageability on verbal tasks in aphasic and deep dysphasic patients have respectively been reported by Bird et al. (2002), Martin and Saffran (1999) and Majerus et al. (2001). Also, in a case study by Baddeley (1993), patient SR was able to learn some English-Finnish word pairs by means of very elaborate semantic associations. Given our results, it is reasonable to assume that patients such as SR or BS, could successfully use readily available visual information when they have to perform a difficult word learning task.

In conclusion, the present work has shown that word imageability, a variable that has been overlooked in previous studies (Baddeley, 1993; e.g. Papagno et al., 1991; Papagno & Vallar, 1992, 1995), determines the degree of verbal WM involvement in paired associate learning of familiar words. Additionally, the current studies demonstrated that the amount of verbal WM resources used to learn associations between familiar and new ('foreign') words is determined by the availability of visual information. Although verbal WM is important in language learning, the locus of its involvement is limited to the learning of phonological representations; the learning of the word associations themselves can rely on other (visual) WM resources.

REFERENCES

Altarriba, J., & Mathis, K. M. (1997). Conceptual and Lexical Development in Second Language Acquisition. *Journal of Memory and Language, 36*(4), 550-568.

Andrade, J., Kemps, E., Werniers, Y., May, J., & Szmalec, A. (2002). Insensitivity of Visual Short-Term Memory to Irrelevant Visual Information. *Quarterly Journal of Experimental Psychology, 55A*(3), 753-774.

Baayen, R., Piepenbrock, R., & Van Rijn, H. (1993). *The CELEX lexical database*. [CD-ROM]. Philadelphia: University of Pennsylvania: Linguistic Data Consortium.

Baddeley, A. (1986). *Working Memory.* Oxford, England: Oxford University Press.

Baddeley, A. (1993). Short-Term Phonological Memory and Long-Term Learning - a Single-Case Study. *European Journal of Cognitive Psychology, 5*(2), 129-148.

Baddeley, A. (1997). *Human Memory: Theory and Practice*. (2nd ed.). Boston: Allyn & Bacon.

Baddeley, A. (2000). The Episodic Buffer: a New Component of Working Memory? *Trends in Cognitive Sciences, 4*(11), 417-423.

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The Phonological Loop as a Language Learning Device. *Psychological Review, 105*(1), 158-173.

Baddeley, A., & Hitch, G. (1974). Working Memory. In Bower, G. A. (Ed.), *Recent Advances in Learning and Motivation.* (pp. 47-90). New York: Academic Press.

Baddeley, A., Papagno, C., & Vallar, G. (1988). When Long-Term Learning Depends on Short-Term Storage. *Journal of Memory and Language, 27*(5), 586-595.

Bird, H., Franklin, S., & Howard, D. (2002). 'little Words' - Not Really: Function and Content Words in Normal and Aphasic Speech. *Journal of Neurolinguistics, 15*(3-5), 209-237.

Bourassa, D., & Besner, D. (1994). Semantics at a Glance - Is the Right-Hemisphere Special. *Canadian Journal of Experimental Psychology-Revue Canadienne De Psychologie Experimentale, 48*(4), 543-552.

Bowey, J. A. (2001). Nonword Repetition and Young Children's Receptive Vocabulary: a Longitudinal Study. *Applied Psycholinguistics, 22*(3), 441-469.

Cheung, H. (1996). Nonword Span as a Unique Predictor of Second-Language Vocabulary Learning. *Developmental Psychology, 32*(5), 867-873.

Coltheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology, 33A*, 497-505.

Duyck W., & Brysbaert M. (2004). Forward and Backward Number Translation Requires Conceptual Mediation Both in Balanced and Unbalanced Bilinguals. Manuscript submitted for publication.

Farmer, E. W., Berman, J. V. F., & Fletcher, Y. L. (1986). Evidence for a Visuospatial Scratch-Pad in Working Memory. *Quarterly Journal of Experimental Psychology Section, 38A*(4), 675-688.

Fearnley, S. (1997). MRC Psycholinguistic Database Search Program. *Behavior Research Methods Instruments & Computers, 29*(2), 291-295.

Freedman, M. L., & Martin, R. C. (2001). Dissociable Components of Short-Term Memory and Their Relation to Long-Term Learning. *Cognitive Neuropsychology, 18*(3), 193-226.

Gathercole, S. E., & Adams, A. M. (1993). Phonological Working-Memory in Very Young-Children. *Developmental Psychology, 29* (4), 770-778.

Gathercole, S. E., & Adams, A. M. (1994). Childrens Phonological Working-Memory - Contributions of Long- Term Knowledge and Rehearsal. *Journal of Memory and Language, 33*(5), 672-688.

Gathercole, S. E., & Baddeley, A. D. (1990). The Role of Phonological Memory in Vocabulary Acquisition - a Study of Young-Children Learning New Names. *British Journal of Psychology, 81*, 439-454.

Gathercole, S. E., Frankish, C. R., Pickering, S. J., & Peaker, S. (1999). 'phonotactic Influences on Short-Term Memory (Vol 25, Pg 84, 1999). *Journal of Experimental Psychology: Learning, Memory and Cognition, 25*(3), 562

Gathercole, S. E., & Hitch, G. J. (1993). Developmental Changes in Short-Term Memory: A Revised Working Memory Perspective. In Collins, A., Gathercole, S. E., Conway, M. A., and Morris, P. E. (Eds.), *Theories Of Memory.* (pp. 189-210). Hove, England: Erlbaum.

Gathercole, S. E., Hitch, G. J., Service, E., & Martin, A. J. (1997). Phonological Short-Term Memory and New Word Learning in Children. *Developmental Psychology, 33*(6), 966-979.

Gathercole, S. E., Service, E., Hitch, G. J., Adams, A. M., & Martin, A. J. (1999). Phonological Short-Term Memory and Vocabulary Development: Further Evidence on the Nature of the Relationship. *Applied Cognitive Psychology, 13*(1), 65-77.

Gathercole, S. E., Willis, C. S., Emslie, H., & Baddeley, A. D. (1992). Phonological Memory and Vocabulary Development During the Early School Years - a Longitudinal-Study. *Developmental Psychology, 28*(5), 887-898.

Gleitman, L. (1993). The Structural Sources of Verb Meanings. In P. Bloom (Ed.), *Core Readings In Language Acquisitions.* (pp. 174-221). Cambridge, England: Harvester Wheatsheaf.

Grasby, P. M., Frith, C. D., Friston, K. J., Bench, C., Frackowiak, R. S. J., & Dolan, R. J. (1993). Functional Mapping of Brain-Areas Implicated in Auditory - Verbal Memory Function. *Brain, 116*, 1-20.

Hanten, G., & Martin, R. C. (2001). A Developmental Phonological Short-Term Memory Deficit: a Case Study. *Brain and Cognition, 45*(2), 164-188.

Henry, L. A., & Millar, S. (1991). Memory Span Increase With Age - a Test of 2 Hypotheses. *Journal of Experimental Child Psychology, 51*(3), 459-484.

Henry, L. A., & Millar, S. (1993). Why Does Memory Span Improve With Age - a Review of the Evidence for 2 Current Hypotheses . *European Journal of Cognitive Psychology, 5*(3), 241-287.

Hulme, C., Maughan, S., & Brown, G. D. A. (1991). Memory for Familiar and Unfamiliar Words - Evidence for a Long- Term-Memory Contribution to Short-Term-Memory Span. *Journal of Memory and Language, 30*(6), 685-701.

Kemps, E., De Rammelaere, S., & Desmet, T. (2000). The Development of Working Memory: Exploring the Complementarity of Two Models. *Journal of Experimental Child Psychology, 77*(2), 89-109.

Kroll, J. F., & Stewart, E. (1994). Category Interference in Translation and Picture Naming - Evidence for Asymmetric Connections Between Bilingual Memory Representations. *Journal of Memory and Language, 33*(2), 149-174.

Logie, R. H. (1986). Visuo-Spatial Processing in Working Memory. *Quarterly Journal of Experimental Psychology, 38A*(2), 229-247.

Majerus, S., Lekeu, F., Van Der Linden, M., & Salmon, E. (2001). Deep Dysphasia: Further Evidence on the Relationship Between Phonological Short-Term Memory and Language Processing Impairments. *Cognitive Neuropsychology, 18*(5), 385-410.

Markman, E. M. (1994). Constraints on Word Meaning in Early Language-Acquisition. *Lingua, 92*(1-4), 199-227.

Martin, N., & Saffran, E. M. (1999). Effects of Word Processing and Short-Term Memory Deficits on Verbal Learning: Evidence From Aphasia. *International Journal of Psychology, 34*(5-6), 339-346.

Martin, R. C. (1993). Short-Term-Memory and Sentence Processing - Evidence From Neuropsychology. *Memory & Cognition, 21*(2), 176-183.

Martin, R. C., Shelton, J. R., & Yaffee, L. S. (1994). Language Processing and Working-Memory - Neuropsychological Evidence for Separate Phonological and Semantic Capacities. *Journal of Memory and Language, 33*(1), 83-111.

Masoura, E. V., & Gathercole, S. E. (1999). Phonological Short-Term Memory and Foreign Language Learning. *International Journal of Psychology, 34*(5-6), 383-388.

Michas, I. C., & Henry, L. A. (1994). The Link Between Phonological Memory and Vocabulary Acquisition. *British Journal of Developmental Psychology, 12*, 147-163.

Nagy, W. E., & Anderson, R. C. (1984). How Many Words Are There in Printed School English. *Reading Research Quarterly, 19*(3), 304-330.

Nagy, W. E., & Herman, P. A. (1987). Breadth and Depth of Vocabulary Knowledge: Implications for Acquisition and Instruction. In McKeown, M. G. and Curtis, M. E. (Eds.), *The Nature of Vocabulary Acquisition.* (pp. 19-35). Hillsdale, NJ: Erlbaum.

Paivio, A., & Smythe, P. C. (1971). Word Imagery, Frequency, And Meaningdulness In Short-Term Memory. *Psychonomic Science, 22*, 333-335.

Papagno, C., Valentine, T., & Baddeley, A. (1991). Phonological Short-Term-Memory and Foreign-Language Vocabulary Learning. *Journal of Memory and Language, 30*(3), 331-347.

Papagno, C., & Vallar, G. (1992). Phonological Short-Term-Memory and the Learning of Novel Words - the Effect of Phonological Similarity and Item Length. *Quarterly Journal of Experimental Psychology, 44A*(1), 47-67.

Papagno, C., & Vallar, G. (1995). Verbal Short-Term-Memory and Vocabulary Learning in Polyglots. *Quarterly Journal of Experimental Psychology, 48A*(1), 98-107.

Paulesu, E., Frith, C. D., & Frackowiak, R. S. J. (1993). The Neural Correlates of the Verbal Component of Working Memory. *Nature, 362*(6418), 342-345.

Poirier, M., & Saintaubin, J. (1995). Memory for Related and Unrelated Words - Further Evidence on the Influence of Semantic Factors in Immediate Serial-Recall. *Quarterly Journal of Experimental Psychology, 48A*(2), 384-404.

Quinn, J. G., & McConnell, J. (1996). Irrelevant Pictures in Visual Working Memory. *Quarterly Journal of Experimental Psychology, 49A*(1), 200-215.

Service, E. (1992). Phonology, Working Memory, and Foreign-Language Learning. *Quarterly Journal of Experimental Psychology, 45A*(1), 21-50.

Sternberg, R. (1987). Most Vocabulary Is Learned From Context. In McKeown, M. G. and Curtis, M. E. (Eds.), *The Nature of Vocabulary Acquisition.* (pp. 89-106). Hillsdale, NJ: Erlbaum.

Van Loon-Vervoorn, W.A. (1985). *Voorstelbaarheidswaarden Van Nederlandse Woorden: 4600 Substantieven, 1000 Verba, 500 Adjectieven*. Lisse: Swets & Zeitlinger.

Walker, I., & Hulme, C. (1999). Concrete Words Are Easier to Recall Than Abstract Words: Evidence for a Semantic Contribution to Short-Term Serial Recall. *Journal of Experimental Psychology: Learning, Memory and Cognition, 25*(5), 1256-1271.

Wetherick, N. E. (1975). Role of Semantic Information in Short-Term-Memory. *Journal of Verbal Learning and Verbal Behavior, 14*(5), 471-480.

Wilson, M. (1988). MRC Psycholinguistic Database - Machine-Usable Dictionary, Version 2.00. *Behavior Research Methods Instruments & Computers, 20*(1), 6-10.

APPENDIX A

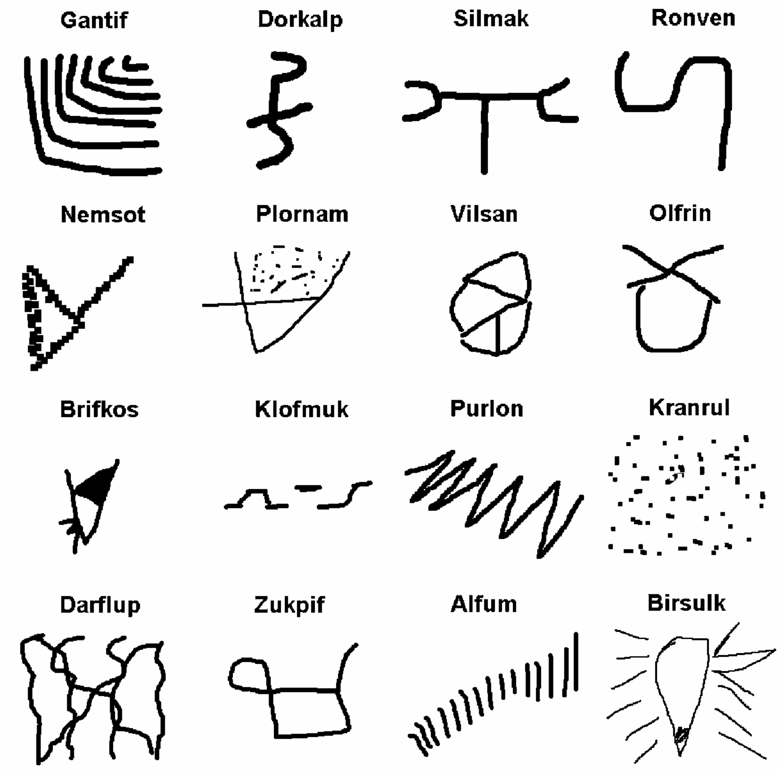
Stimuli Experiment 1 and 3 (English translations between brackets)

Word Type	List One	List Two
Word-Concrete Word Pairs (Experiments 1 and 3)	auto [car]-armband [bracelet]	balpen [ballpoint]-geweer [gun]
	badpak [bathing suit]-ladder [ladder]	kamer [room]-aardbei [strawberry]
	bliksem [lightening]-vlinder [butterfly]	konijn [rabbit]-mummie [mummy]
	sleutel [key]- matras [mattress]	nagel [nail]-paleis [palace]
	tafel [table]-oorbel [earring]	oven [oven]-rugzak [rucksack]
	viool [violin]-kasteel [castle]	parfum [perfume]-tractor [tractor]
	kikker [frog]-spijker [nail]*	bontjas [fur coat]-jongen [boy]*
	vinger [finger]-augurk [gherkin]*	koelkast [refrigerator]-lippen [lips]*
Word-Abstract Word Pairs (Experiment 1)	auto [car]-verzoek [request]	balpen [ballpoint]-inteelt [inbreeding]
	badpak [bathing suit]-geding [lawsuit]	bontjas [fur coat]-noodlot [fate]
	bliksem [lightening]-beschik [disposal]	kamer [room]-subject [subject]
	kikker [frog]-toeval [coincidence]	koelkast [refrigerator]-profijt [profit]
	sleutel [key]- tactiek [tactics]	konijn [rabbit]-welzijn [wellbeing]
	tafel [table]-schennis [violation]	nagel [nail]-talent [talent]
	vinger [finger]-stemming [mood]	oven [oven]-schande [shame]
	viool [violin]-voorval [incident]	parfum [perfume]-bijnaam [nickname]
Word-Function Word Pairs (Experiment 3)	auto [car]-sedert [since]	balpen [ballpoint]-vanaf [from]
	badpak [bathing suit]-ofwel [either]	kamer [room]-tenzij [unless]
	bliksem [lightening]-terwijl [while]	konijn [rabbit]-indien [if]
	sleutel [key]- zodat [so (that)]	nagel [nail]-omdat [because]
	tafel [table]-daarom [therefore]	oven [oven]-wegens [due to]
	viool [violin]-misschien [perhaps]	parfum [perfume]-wanneer [when]
Word-Nonword Pairs (Experiments 1 and 3)	auto [car]-plornam	balpen [ballpoint]-alfum
	badpak [bathing suit]-vilsan	kamer [room]-kranrul
	bliksem [lightening]-olfrin	konijn [rabbit]-brifkos
	sleutel [key]-ronven	nagel [nail]-zukpif
	tafel [table]-dorkalp	oven [oven]-purlon
	viool [violin]-silmak	parfum [perfume]-klofmuk
	kikker [frog]-nemsot*	bontjas [fur coat]-birsulk*
	vinger [finger]-gantif*	koelkast [refrigerator]-darflup*

* These stimuli were not used in Experiment 3.

APPENDIX B

The sixteen line drawings with their corresponding nonwords used in Experiment 2.



CHAPTER 4
TRANSLATION AND ASSOCIATIVE PRIMING WITH CROSS-LINGUAL PSEUDOHOMOPHONES: EVIDENCE FROM DUTCH-ENGLISH BILINGUALS.

Manuscript submitted for publication[1, 2]

Using a masked priming paradigm with a lexical decision task performed by Dutch-English bilinguals, we showed that the recognition of visually presented L1 (e.g. *TOUW*) and L2 (e.g. *BACK*) targets is facilitated by respectively L2 and L1 primes, which are pseudohomophones (*roap* and *ruch*) of the target's translation equivalent (*rope* and *rug*). In two further experiments, we found that recognition of L2 targets (e.g. *CHURCH*) was also facilitated by L1 pseudohomophones (e.g. *pous*) of related words (paus [pope]). Contrastingly, no significant effect was obtained for L1 targets (e.g. *BEEN* [leg]) and L2 pseudohomophone associative primes (e.g. *knea*). In two last experiments, we found that a L2 target word (e.g. *CORNER*) is facilitated by an L2 (intra-lingual) homophone (e.g. *hook*) of its L1 translation equivalent (hoek). The same was not true for respective L1 targets (e.g. *DAG* [day]) and primes (e.g. *dij*) These findings are in line with recent research on language-independent activation of phonological representations in bilinguals (Brysbaert, Van Dyck, & Van De Poel, 1999; Van Wijnendaele & Brysbaert, 2002).

---

[1] This paper was co-authored by Marc Brysbaert

[2] The authors are indebted to Dominiek Sandra for several suggestions on the interpretation of our data and to Wouter De Baene and Timothy Desmet for excellent comments on an earlier version of this paper.

INTRODUCTION

Ever since the recent development of literature on bilingual language processing, it has intrigued researchers whether languages mastered by bilinguals are processed by functionally and structurally independent systems or not. However, whereas words are represented through at least three different representational levels (i.e. in an orthographic lexicon, a phonological lexicon and semantically), research has mainly focused on lexical[3] autonomy ('the mental dictionary'). During the last decade, there has been a lively debate whether lexical access during visual word recognition is language-specific or whether lexical representations of both languages interact early during this process. Now, it seems that this debate has almost been settled in favor of the latter hypothesis (for a recent review, see Dijkstra & Van Heuven, 2002). Also, there has been some discussion whether both languages activate semantic representations to the same extent. Whereas the mainstream hypothesis (e.g. Kroll & Stewart, 1994) suggests that second language (L2) lexical representations only indirectly activate semantics through their first language (L1) translations, recent studies have found strong indications of strong L2 lexico-semantic links for certain types of words (e.g. Duyck & Brysbaert, 2002; Duyck & Brysbaert, 2004; Francis, Augustini, & Saenz, 2003).

As for phonological representations, bilingual research is much scarcer, in contrast with the large body of evidence focusing at phonological coding in the monolingual domain. There have only been a few studies (e.g. Brysbaert, Van Dyck, & Van De Poel, 1999; Jared & Kroll, 2001; Van Wijnendaele &

---

[3] In agreement with the literature on bilingualism (e.g. Dijkstra & Van Heuven, 2002), we will use term 'lexical' with reference to orthographic representations. Entries in the phonological lexicon will be labelled as 'phonological representations'.

Brysbaert, 2002) which were directly aimed at investigating whether the early activation of phonological representations during word recognition is language-independent (just as for lexical representations) or not. The present study was set up to find additional evidence with respect to this issue by integrating the effects typically found in studies focusing on the three representational levels mentioned above. Because the experiments in this study are based on the masked phonological priming paradigm, we will first briefly discuss relevant earlier studies in the literature on respectively monolingual and bilingual visual word recognition.

PHONOLOGICAL CODING IN MONOLINGUAL VISUAL WORD RECOGNITION

Whereas the classical dual-route model of visual word recognition (e.g. Coltheart, 1978; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001) attached quite some importance to a direct route from orthography to meaning (at least for skilled readers), recent accounts of visual word recognition have particularly stressed the importance of (early and automatic) phonological coding (e.g. Berent & Perfetti, 1995; Frost, 1998). Most of the evidence for the latter claim comes from the masked phonological priming paradigm, first used by Humphreys, Evett and Taylor (1982). Using a perceptual identification task, they found that tachistoscopically presented words (e.g. *made*) were easier to be recognized when preceded by a homophone prime (e.g. *maid*) than by a graphemic control word (e.g. *mark*). However, while being the first to show that phonological codes are accessed so early during the word recognition process, their results did not allow to make very strong claims about this issue, as they did not succeed in replicating this finding with pseudohomophone primes. This suggests that at least some of their phonological (word) priming effect was of a lexical nature, and perhaps not due to the activation of phonological representations. It was almost ten years later that Perfetti and Bell (1991) succeeded in passing this more convincing pseudohomophone test (using slightly longer prime durations), an effect later replicated by Grainger and Ferrand (1996). Also, this pseudohomophone effect has been extended to the lexical decision task (Ferrand & Grainger, 1992; Ferrand & Grainger, 1993; Frost, Ahissar, Gotesman, & Tayeb, 2003; Grainger & Ferrand, 1996; Lukatela, Frost, & Turvey, 1998) and to naming (Lukatela & Turvey, 1994b; see Kim & Davis, 2002, for similar results in a Korean naming task using word primes written in a different script).

Importantly for the present study, Lesch and Pollatsek (1993) extended the previously mentioned effects and showed that the naming of a target word (e.g. *nut*) was not only facilitated by a "real" semantic associate (e.g. *beech*) prime, but also by a homophone of that associate (e.g. *beach*), relative to a graphemic control word (e.g. *bench*). This was replicated by Lukatela and Turvey (1994a) with nonword primes (i.e. naming *frog* was facilitated by *tode*, a pseudohomophone of *toad*). In their view, the associative prime or its (pseudo)homophone activates the shared phonological representation. This feeds back to the corresponding semantic or lexical representation (for a discussion of a lexical-associative vs. semantic locus of priming, see the *General Discussion*), which in turn activates the (lexical or semantic) representations of related words. Among these is the target word, which is thus preactivated, leading to faster responses. Note that this pseudohomophone associative priming effect has recently been replicated and extended to the lexical decision task by Drieghe and Brysbaert (2002). Also, Tan and Perfetti (1997) found that a Chinese target can be primed with a homophone of a target synonym (basically an almost maximally associated word), even though phonological coding of Chinese orthography is much less straightforward and efficient (but see Zhou & Marslen-Wilson, 1999). Finally, whereas the associative pseudohomophone priming effect arises from activation flowing from phonology to semantic/lexical representations, there are also strong indications that semantic representations activated by a prime can activate phonology. Farrar, Van Orden and Hamouz (2001) for instance, found that the word prime *sofa* inhibits the recognition *touch*, through activation of its semantic associate *couch*, which is body-rime inconsistent with the target.

As a conclusion, these findings from the monolingual domain strongly suggest that (a) phonological representations are activated automatically and very early (pre-lexically) during visual word recognition (e.g. Frost, 1998; e.g. Grainger & Ferrand, 1996; Lukatela & Turvey, 1994b; Perfetti & Bell, 1991; see Zhou & Marslen-Wilson, 1999, for a discussion of non-alphabetic languages, such as Chinese, which have less transparent grapheme-phoneme conversion rules) and (b) the activation spreading from these phonological representations is fast and strong enough to pre-activate semantic/lexical (see *General Discussion*) representations and therefore influences the recognition of associated words (e.g. Lesch & Pollatsek, 1993; Lukatela & Turvey, 1994a; Drieghe & Brysbaert, 2002).

ACTIVATION OF PHONOLOGICAL REPRESENTATIONS IN BILINGUALS

In the present section, we will briefly discuss to what extent the previous findings apply to bilingual visual word recognition. As mentioned earlier, there are only a few studies directly assessing this issue. Among the first was that of Brysbaert, Van Dyck and Van de Poel (1999). They reasoned that it is very likely, given the evidence mentioned above, that the (not strategically controllable) phonological coding of visually presented words occurs for all grapheme-to-phoneme conversion (GPC) rules mastered by a bilingual person. Moreover, given its timecourse, this probably occurs before a language selection system (if any, see the discussion on language-independent lexical activation, Dijkstra & Van Heuven, 2002) gets involved. Hence, it was expected that early phonological coding through L1 GPC rules occurs during visual word recognition in L2 and vice versa.

The first part of this statement was examined in the Brysbaert et al. study itself, using a perceptual identification task with French monolinguals and Dutch-French bilinguals. First, they replicated the pseudohomophone priming effect of Grainger and Ferrand (1996, see earlier). Second, they replicated this effect with Dutch-French bilinguals, using the same French (L2 for these participants) stimuli, which shows that the same processes underlie L1 and L2 word recognition. Third and most importantly, these bilingual participants (but not the monolinguals) also showed a similar cross-lingual phonological priming effect. It was easier to identify L2 targets (e.g. *nez* [nose]) following L1 homophonic primes (e.g. *nee* [no]) than those following L1 graphemic control primes (e.g. *nek* [neck]). This effect was replicated more recently by Duyck, Diependaele, Drieghe and Brysbaert (2004). Note that the L1 homophonic primes were only homophones of the L2 target word according to L1 (Dutch) grapheme-to-phoneme conversion rules, which shows that L1 GPC rules were processed even though the participants were performing a task in their L2. Finally, it is important to note that Brysbaert et al. also replicated their effect with nonword primes (i.e. the L1 pseudohomophone *poer* primed the L2 target *pour* relative to the graphemic control prime *poir*) (the pseudohomophone test, see earlier), which strongly argues against a lexical locus for this cross-lingual phonological priming effect. The above findings are compatible with Gollan, Forster and Frost (1997), who investigated translation priming with Hebrew-English bilinguals in a lexical decision task. They reported that the priming effect of L1 translation primes on the recognition of L2 targets was larger if both words were phonologically similar (an interaction effect also found in a naming task by Kim & Davis, 2003, Experiment 2). Also, Dijkstra, Grainger and Van Heuven (1999) showed that lexical decision RTs are larger for L2 words which are phonologically similar to L1 words, even though L1 phonology was not useful for this task. Further evidence for automatic phonological coding of L1 words during L2 word processing comes from a Korean-English study of Kim and Davis (2003). They found that L2 targets are primed by L1 homophone primes, even though both languages have different alphabets, both in a naming task and a lexical decision task (although the latter effect was only significant in a one-tailed test, and not acknowledged by the authors).

Some years later, evidence for the second (stronger) part of the statement mentioned above was obtained: reversing the language dominance of the participants while using the same pseudohomophones as Brysbaet al. (1999), Van Wijnendaele and Brysbaert (2002) showed that recognition of L1 targets is facilitated by L2 homophonic primes (relative to graphemic control primes). Hence, because the prime is only a pseudohomophone of the target according to Dutch (now L2) GPC rules, this strongly suggests that pre-lexical phonological coding during visual word recognition also occurs through L2 GPC rules, even when performing a task in the native language. This is compatible with a much earlier finding of Nas (1983), who showed that Dutch-English bilinguals are slower to reject nonwords (e.g. *snay*) in a Dutch (L1) lexical decision task which sound like (according to L2 GPC rules) existing Dutch words (e.g. *snee* [cut]). However, one might object that the activation of L2 phonological representations in the latter study might have been triggered by the presence of typical L2 (English) letter combinations (e.g. *ay* in the example above) which are illegal in Dutch (L1). Finally, it is important to note that the just mentioned interaction of the L1-L2 translation priming effect of Gollan et al. (1997, see earlier) with the phonological overlap between prime and target was not present if the stimulus languages were switched (L2 primes/L1 targets). This suggests that the phonological coding of the L2 prime did not occur or was at least not strong enough to influence L1 word recognition.

While the two previously discussed studies (Brysbaert et al., 1999; Van Wijnendaele & Brysbaert, 2002) unambiguously point to language-independent, simultaneous activation of phonological representations (even those of the non-active language), the study of Jared and Kroll (2001) led to more differentiating results. They found that French-English bilinguals were slower to name L2 words which have word-body enemies in L1 (e.g. the English word *bait* contains the word body *ait* which is pronounced differently in French) relative to controls (e.g. the word *bump* contains the letter sequence *ump* which is illegal in French). Later, Jared and Szucs (2002) found similar results for interlingual homographs which have conflicting pronunciations in English and French (e.g. *pain*). Like the Brysbaert et al. (1999) study, this suggests that L1 phonological coding is automatically engaged, even when performing a task in L2. However, in contrast to Van Wijnendaele and Brysbaert (2002), Jared and Kroll did not find straightforward evidence for the opposite. English-French bilinguals were slower to name the same English words (now L1) having French (L2) enemies relative to controls, but only after they had just named a set of French filler words (see also Jared et al., 2002). Hence, activation of L2 phonological representations was only present if L2 GPC rules had been active just before the L1 task.

In conclusion, these few studies on language-independent activation of phonological representations strongly suggest that visually presented words are always automatically processed through L1 GPC rules, even when reading in L2. Evidence for the opposite statement is mixed. Whereas the results of Van Wijnendaele and Brysbaert (2002) clearly show that L2 phonological representations are accessed during L1 word recognition, findings of Jared and Kroll (2001) suggest that the activation in these representations may only be strong enough to influence L1 processing if L2 GPC rules have recently been active.

ACTIVATION OF LEXICAL REPRESENTATIONS IN BILINGUALS

Whereas the primary focus of this paper is on bilingual phonological coding, it is also important to discuss very briefly the current state of affairs with respect to the activation of lexical knowledge, before going into detail about the present experiments. This will appear to be important when discussing the motivation and results of the current study. As noted in the beginning of this paper, there is a growing consensus that visually presented words activate lexical representations from both languages in bilinguals. Evidence for this statement comes from studies in which L1 lexical knowledge, though irrelevant to the task at hand, influences L2 language processing. For instance, Dijkstra, Timmermans and Schriefers (2000) showed that Dutch-English bilinguals respond slower to interlingual homographs (i.e. words which exist in both L1 and L2, but have a different meaning, e.g. *room* means cream in Dutch) than to words which only exist in L2 in a lexical go/no-go task (press a button only if the target is a word in L2). Moreover, Van Hell and Dijkstra (2002) recently showed that L2 and even L3 lexical knowledge also influences L1 lexical access in an exclusive native language context. They reported faster lexical decision responses of Dutch – English – French trilinguals for L1 targets having L2 and L3 near-cognate (i.e. orthographically nearly identical) translation equivalents (e.g. *brood – bread*) than for control words. This shows that L2 (and even L3) lexical representations are accessed during L1 word recognition and that their activation is strong enough to influence L1 representations. Because this is beyond the scope of this paper, we refer to Dijkstra and Van Heuven (2002) for a comprehensive and recent review of further evidence in favor of non-selective lexical access.

THE PRESENT STUDY

The present study was set up to find additional evidence for the claim of pre-lexical language-independent activation of phonological representations made by Brysbaert et al. (1999) and Van Wijnendaele and Brysbaert (2002). As we will discuss in this section, this was done by subsequently extending the previously discussed monolingual (a) pseudohomophone priming effect and (b) pseudohomophone associative priming effect to a bilingual context.

The first two experiments that we will report constitute a cross-lingual extension of the pseudohomophone priming effect discussed earlier (e.g. Perfetti & Bell, 1991; Lukatela & Turvey, 1994b; Lukatela et al., 1998). In these monolingual studies, it was shown that recognition of a target word (e.g. *TOAD*) is facilitated when it is preceded by a pseudohomophone prime (e.g. *tode*). In this study, we will explore whether it is also possible to facilitate the recognition of a L2 target word (e.g. *BACK*) with a L1 pseudohomophone (e.g. *ruch*) of its L1 translation equivalent (rug) (relative to a graphemic control nonword prime which shares the same letters with the target, e.g. *gect*). If L1 phonological representations are indeed pre-lexically assembled during L2 target recognition (Brysbaert et al., 1999), the L1 pseudohomophone should quickly activate its phonological representation (identical to that of the "real" translation equivalent), which then activates the corresponding semantic representation. Given the evidence for non-selective lexical access (e.g. Dijkstra & Van Heuven, 2002), this should trigger pre-activation of the corresponding L2 target, in the same way activation spreads to related intra-lingual lexical entries in pseudohomophone associative priming (see earlier, Lukatela & Turvey, 1994a). This was investigated in Experiment 1. If L2 GPC rules are also active during L1 processing (Van Wijnendaele & Brysbaert, 2002), the same line of reasoning should apply to L1 targets (e.g. *TOUW*) and L2 pseudohomophone (e.g. *roap*) primes (control prime *joll*). This was investigated in Experiment 2. In these two experiments, we also manipulated cognate status (i.e. cross-linguistic form overlap) of the two translation equivalents involved. For instance, Experiment 1 included non-cognate stimuli (e.g. *ruch* [rug] – BACK), as well as cognate stimuli (e.g. *oleif* [olijf] – OLIVE). Note that these two experiments are in fact very similar to those of Tan and Perfetti (1997, see earlier), who reported that L1 target recognition is facilitated by a prime which is homophone to a synonym of the target. In these experiments, however, the synonym is a translation equivalent. According to the strong non-selective lexical access view, these are both basically cases of different lexical labels representing the same meaning.

The second part of this study was designed as a cross-lingual version of the pseudohomophone associative priming effect (see earlier, Lukatela & Turvey, 1994a; Drieghe & Brysbaert, 2002). In the monolingual effect, target recognition (e.g. *FROG*) is facilitated by a pseudohomophone (e.g. *tode*) of a related word (toad). In Experiment 3, we will explore whether it is possible to prime a L2 target (e.g. *CHURCH*) with a L1 pseudohomophone (e.g. *pous*) of a related L1 word (paus [pope]) (graphemic control prime: *zeun*). Unlike the first two experiments, the prime is now a pseudohomophone, not of its translation equivalent, but of a word related to it. The same line of reasoning applies: the L1 pseudohomophone *pous* (homophone of *paus*, meaning *pope*) should activate its phonological representation, which in turn activates the semantic representation of *paus*. This activation is then spread to related semantic representations (Lukatela & Turvey, 1994a), like that of *kerk*, which leads to faster responses to the L2 word (*CHURCH*) representing that meaning. Again, we also carried out the same experiment with L1 targets (e.g. *BEEN* [leg]) and L2 pseudohomophone associative primes (e.g. *knea*) (Experiment 4).

Finally, as the third part of this study, we also carried out two experiments in which we investigated whether it is possible to prime target words with intra-lingual homophones of their translation equivalents. For instance, in Experiment 5, we explored whether recognition of L2 words such as *CORNER* is facilitated by L2 primes such as *hook*, which is a homophone of the Dutch (L1) word *hoek* (meaning *corner*). This would offer evidence for (a) pre-lexical phonological coding of L2 primes and (b) for language-independent semantic activation of phonological representations (i.e. the phonological representation associated with *hook*, /huk/, activates both its L1 and L2 meaning). In Experiment 6, this was replicated for L1 (e.g. *dij* [thigh] – DAG [day]).

EXPERIMENT 1

METHOD

Participants. The participants were 22 Dutch-English bilinguals. Mean age was 20.8 years (SD = 4.22). They were all students at Ghent University, participating for course requirements. They had started to learn English in a scholastic setting around the age of 14-15. All participants lived in a L1 dominant environment, speaking Dutch at home, at school, with friends, etc. All of them were regularly exposed to their L2 (English) (music, internet, films, television, etc.).

Stimulus Materials. The stimuli consisted of 56 L2 (English) word targets and 56 L2 nonword targets. Half of the word targets were words for which the corresponding L1 (Dutch) translation equivalent is a (near) cognate (e.g. *sand* [zand]). The other half of the word targets were non-cognates (e.g. *bucket* [emmer]). All word targets were matched with two types of L1 (Dutch) nonword primes (see Appendix A). The first type of primes were pseudohomophone translation primes, i.e. L1 (Dutch) nonwords which have the same pronunciation as the L1 translation equivalent of the L2 target (e.g. *ruch* [rug] – BACK). The second type of primes were graphemic control primes, i.e. L1 nonwords which have the same letters in common with the L2 target as the L1 pseudohomophone translation prime (e.g. *gect* – BACK), for the stimulus pair mentioned above, in which the prime shares a letter *c* in the third letter position). This constraint was set to ensure that any priming effect is not due to orthographic overlap of the pseudohomophone prime with the target word. These control primes always had the same number of letters as the corresponding pseudohomophone prime. In addition, we also made sure that both types of nonword primes were equally wordlike. If this were not the case, it could be argued that pseudohomophone nonword primes are intrinsically more wordlike than other random nonwords, because they only contain legal grapheme-to-phoneme conversion rules (otherwise they would not be pronounced as real words). Therefore, it is possible that they would trigger a 'word' response to a larger degree than less wordlike nonwords, causing faster responses to the following (word) targets. To control for wordlikeness, we matched the two types of primes with respect to two variables, i.e. summated bigram frequency and neighbourhood density. The first refers to the summated number of occurrences of each of the nonword's bigrams (e.g. the nonword *gect* contains three bigrams, i.e. *ge*, *ec* and *ct*) in the CELEX lexical database (Baayen, Piepenbrock, & Van Rijn, 1993). The latter variable refers to the number of orthographic neighbours (i.e. words that have all but one letter in common with a given nonword) within that CELEX database for that language. It is very plausible to assume that wordlike nonwords contain bigrams which are more frequent in a given language and have more neighbours than less wordlike nonwords. This matching was done by means of the WordGen program (Duyck, Desmet, Verbeke, & Brysbaert, 2004), which uses the CELEX database to generate wordlike nonwords satisfying different combinations of bigram frequency and neighbourhood density constraints. The program was probed for a nonword within close range of the respective values for each of the to-be-matched pseudohomophone. The program was also set to exclude bigrams which never occur (as the onset, suffix, or any other part of the word) in L1 (Dutch). Using this procedure, two sets of pseudohomophones and pronounceable (in L1) graphemic control primes were obtained, which were matched for summated bigram frequency (cognates: respectively $M = 24240$ and $M = 24599$, $F < 1$; non-cognates: $M = 26294$ and $M = 25952$, $F < 1$), neighbourhood size (cognates: $M = 5.36$ and $M = 5.04$, $F < 1$; non-cognates: $M = 6$ and $M = 5.68$, $F < 1$), word length (identical) and orthographic overlap with the target. Care was also taken that no control prime sounded like an existing Dutch or English word. Also, English targets did not sound like, or were homographs of existing Dutch words. Finally, non-cognates and cognates were matched with respect to word frequency (CELEX log frequency per million, $M = 1.70$ and $M = 1.73$, $F < 1$). The 56 L2 nonword targets were each matched to one of the L2 word targets following the procedure of nonword generation mentioned above. Hence, the nonword targets were wordlike, pronounceable (in L2) letter strings, matched with word targets with respect to length and wordlikeness.

To conclude this section, we would like to point out that the pseudohomophone translation primes were matched with the graphemic control primes regarding the number of letters they share with the target (e.g. because the prime *ruch* shares a *c* in the third letter position with *BACK*, so does the control prime *gect*). Because in any alphabetic language, phonological overlap is almost always partially correlated with orthographic overlap, the pseudohomophone prime (in most cases) has some letters in common with the translation equivalent from which it is a pseudohomophone (in this case 'rug'). This is not problematic for any phonological priming effect in the context of the present study for a number of reasons. First, the 'real' translation equivalent was not actually shown during the experiment (unlike the target for which orthographic overlap was controlled). Therefore, any effect of the *ru* which appears in *ruch* and 'rug' would be an indirect effect, in that *ruch* pre-activates a little bit any word starting with an *r*, any word with an *u*, any word starting with *ru*, etc. Among those hundreds of words is also 'rug', which could feed some of the little activation it receives to its translation equivalent *back*. We believe such an indirect effect, mediated by the limited activation in 'rug' is likely to be outweighed by the much stronger phonological manipulation. Second, Perea and Lupker (2003) actually tested whether such indirect activation occurs with nonword primes. Using the same masked priming paradigm, they found that the nonword prime *judpe* did not prime the associated target *court* (three nonsignificant effects: 6 ms, 0 ms and 4 ms), whereas the associated prime *judge* (14ms, 15 ms and 19 ms, all significant), and its transposed internal letter nonword prime *jugde* did (effects of 12 ms, 10 ms and 15 ms all significant). If an indirect orthographic priming mechanism such as described above would be active, on would expect *judpe* to activate *judge* (among other words) to a certain extent (probably much more than *ruch* would pre-activate 'rug', given the larger number of common letters), causing a priming effect. Moreover, if such a process does not elicit such an orthographic effect between monolingual stimulus pairs, it is unlikely to do so across languages. Third, in a similar Spanish study, Carreiras and Perea (2002) found that monosyllabic nonword primes (as most of the primes in this study are) did not prime monosyllabic targets sharing the first two letters (e.g. *blan* – *BLOC*), even though this should be a much more direct and stronger effect than any indirect orthographic priming in this study. Fourth, Lukatela, Savic, Urosevic and Turvey (1997) found that the target the target *robot* (/robot/) was significantly more primed by the mixed-alphabet (Roma-

Cyrillic) but phonologically unique nonword ROБOT (/robot/) than by the phonologically ambiguous nonword prime POBOT (/robot/ or /rovot/ or /pobot/ or /povot/) and even than by the phonologically ambiguous, but orthographically identical word prime ROBOT (/robot/ or /rovot/). Later, Lukatela, Carello, Savic, Urosevic and Turvey (1999) reported similar findings for the same primes in associative priming (target *automat*). Again, this is not compatible with an indirect orthographic priming effect such as described above. Fifth, Kim and Davis (2002) showed that phonological priming occurs in the absence of orthographic overlap, making use of two different alphabets of the Korean writing system. All the above findings make it very likely that any large priming effect found is due to phonology, and not to an indirect, three-step orthographic process.

Procedure. Participants were tested in small groups. Care was taken that they were placed sufficiently far from each other. It was not possible to see the computer screen of another participant. Similar to Drieghe and Brysbaert (2002), participants received written instructions to perform a lexical decision task. This adds further strength to any phonological effect found, as the lexical decision task (e.g. Lukatela & Turvey, 1994a), does not explicitly require access to phonology, unlike naming (see Taft & Van Graan, 1998), which was used by Lukatela and Turvey (1994a) for example. These instructions mentioned that ten practice trials and several experimental trials would follow. No indication was given with regard to the presence of shortly presented words (primes) during the experiment. The participants were instructed to react to the target word and press one button if the presented letter string was an existing English word, or another button if this was not the case. Practice trials were followed by feedback concerning the correctness of the response, whereas no feedback was given after the experimental trials. Each participant completed 112 experimental trials (28 cognate L2 word targets, 28 non-cognate L2 word targets and 56 L2 nonword targets) in a random order. Each of the targets was only presented once. For each of the participants, 28 (14 cognate and 14 non-cognate) of the word targets were presented with a pseudohomophone prime, and 28 were presented with a control prime. Each participant received a different random permutation. Across participants, all target words were presented with each prime.

Each trial started with a forward mask (consisting of six hash-marks, ######) presented for 500 ms. This mask was followed by the presentation of the prime for 57 ms (similar to Lukatela & Turvey, 1994a; Drieghe & Brysbaert, 2002), a backward mask for 57 ms (similar to Lesch & Pollatsek, 1993), and the target. Stimulus presentation was synchronized with the refresh cycle of the screen (70 Hz), using timing routines described by Bovens and Brysbaert (1990). The prime appeared in lowercase letters, unlike the target which was displayed in uppercase letters. The target remained on the screen until the participant gave a response (using a response box connected through the computer's gameport). Throughout the experiment, two vertical lines were displayed centred on the screen, with a gape between them of approximately 1 cm. Participants were instructed to look at the gap between these lines. Both masks and stimuli were presented so that the second character always appeared between these two lines. Earlier studies (e.g. Brysbaert, Vitu, & Schroyens, 1996) have shown that this is the optimal viewing position for short words.

RESULTS

The proportion of false responses to L2 word and nonword targets was 12.3%. This is higher than the accuracy level generally observed in L1 lexical decision tasks (see also *Experiment 2*). These trials were excluded from all analyses. An ANOVA was performed with Cognate Status (cognate vs. non-cognate) and Primetype (graphemic control vs. pseudohomophone) as repeated measures factors. The dependent variable was the mean RT across trials. Mean RTs and proportion of correct trials as a function of these two independent variables are presented in Table 1.

The effect of cognate status tended slightly towards significance, $F_1(1, 21) = 3.20$, $MSE = 4657$, $p < .09$, $F_2(1, 54) = 1.11$, $MSE = 15702$, $p < .30$.

	Non – Cognates			Cognates		
	Example	RT	% errors	Example	RT	% errors
Orthographic Control Prime	*gect* – BACK	675	14.8	*ogt* – EIGHT	706	16.4
Pseudohomophone Translation Prime	*ruch* [rug] – BACK	644	13.6	*agt* [acht] – EIGHT	665	14.9
Net Priming Effect		31	1.2		41	1.5

Table 1. Mean RTs (ms) and accuracy (% errors) as a function of cognate status and primetype (Experiment 1, L1 pseudohomophone translation primes - L2 targets). L1 homophone translation equivalents are displayed between brackets (these words were not presented during the experiment).

Responses to cognate targets (686 ms) were slightly slower than to non-cognate targets (660 ms). Most importantly, the main effect of primetype was significant, both in the analysis by participants and by items, respectively $F_1(1, 21) = 10.61$, $MSE = 2666$, $p < .01$ and $F_2(1, 54) = 5.40$, $MSE = 9026$, $p < .03$. Responses to targets following a pseudohomophone translation prime (655 ms) were significantly faster than those following a graphemic control prime (691 ms). This priming effect did not interact with cognate status (both F*s* < 1), although Table 1 shows a slightly larger priming effect for cognate targets. Planned comparisons showed that the priming effect was significant both in the cognate and non-cognate conditions, respectively $F_1(1, 21) = 5.62$, $MSE = 3296$, $p < .01$, $F_2(1, 54) = 2.96$, $MSE = 9026$, $p < .05$ and $F_1(1, 21) = 3.74$, $MSE = 2770$, $p < .05$, $F_2(1, 54) = 2.45$, $MSE = 9026$, $p < .06$ (because we had well-founded expectations concerning the priming effect at the onset of this study, *p* values are reported for one-tailed tests). There were no effects of cognate status and primetype on the proportion of correct trials (all *F*s < 1), as Table 1 already suggests.

DISCUSSION

In line with expectations, we found a significant forward (L1-L2) pseudohomophone translation priming effect: L2 target words were recognized faster if they were preceded by a L1 pseudohomophone of their L1 translation equivalents. In line with Brysbaert et al. (1999), this shows that the pseudohomophone primes were phonologically coded through L1 GPC rules, even though the task only involved L2 target words. Moreover, these phonological representations were activated strongly enough to pre-activate the underlying semantic representations and corresponding L2 translation equivalents. This effect did not interact with the degree of form overlap (cognate status) between the translation equivalents. For further theoretical implications of these findings, we refer to the *General Discussion.*

EXPERIMENT 2

METHOD

Participants. The participants were 22 Dutch-English bilinguals. Mean age was 21 years (SD = 4.01). None of them participated in Experiment 1. They belonged to the same population and had a similar L2 history as the participants in Experiment 1.

Stimulus Materials. The stimulus list was similar to Experiment 1, but the languages were switched. The stimuli consisted of 56 L1 (Dutch) word targets and 56 L1 nonword targets. Half of the word targets were words for which the corresponding L2 (English) translation equivalents is a (near) cognate (e.g. *melk* [milk]). The other half of the word targets were non-cognates (e.g. *auto* [car]). All word targets were matched with two types of L2 (English) nonword primes (see Appendix B). The first type of primes were pseudohomophone translation primes, i.e. L2 nonwords which have the same pronunciation as the L2 translation equivalent of the L1 target (e.g. *trea* [tree] – BOOM). All these pseudohomophones were drawn from the ARC nonword database, a large set of pseudohomophone letter strings composed following strict criteria described by Rastle, Harrington and Coltheart (2002). The second type of primes were L2 nonword graphemic control primes, matched with the pseudohomophones following the criteria and procedure used in Experiment 1. The resulting two sets of pseudohomophones and pronounceable (in L2) control primes were matched for summated bigram frequency[4] (cognates: respectively $M = 5920$ and $M = 5852$, $F < 1$; non-cognates: $M = 4614$ and $M = 4603$, $F < 1$), neighbourhood size (cognates: $M = 4.32$ and $M = 4.46$, $F < 1$; non-cognates: $M = 5.5$ and $M = 5.54$, $F < 1$), word length (identical) and orthographic overlap with the target. Care was also taken that no control prime sounded like an existing Dutch or English word. Also, Dutch (L1) targets did not sound like, or were homographs of existing English words. Finally, non-cognates and cognates were matched with respect to word frequency (CELEX log frequency per million, $M = 1.62$ and $M = 1.59$, $F < 1$). The 56 L1 nonword targets satisfied the criteria mentioned above for the corresponding stimuli in Experiment 1.

Procedure. The procedure was identical to that of Experiment 1. Again, the presence of (L2) primes in the experiment was not mentioned.

RESULTS

The proportion of false responses was 4.4%. Again, these trials were excluded from all analyses. An ANOVA was performed with Cognate Status (cognate vs. non-cognate) and Primetype (graphemic control vs. pseudohomophone) as repeated measures factors. The dependent variable was the mean RT across trials. Mean RTs and proportion of correct trials as a function of these two variables are presented in Table 2.

Responses to cognate targets (583 ms) were slightly slower than to non-cognate targets (576 ms). This difference was not significant, $F_1(1, 21) =$

1.34, MSE = 812, p > .26, F_2 < 1. Most importantly, the main effect of primetype was significant, just as in Experiment 1 (L2 targets), $F_1(1, 21) = 9.90$, $MSE = 998$, $p < .01$ and $F_2(1, 54) = 969.54$, $MSE = 9511$, $p < .001$.

	Non – Cognates			Cognates		
	Example	RT	% errors	Example	RT	% errors
Orthographic Control Prime	*joll* – TOUW	587	4.0	*preef* – DIEF	594	4.8
Pseudohomophone Translation Prime	*roap* [rope]– TOUW	566	3.7	*theef* [thief] – DIEF	573	3.8
Net Priming Effect		21	0.3		21	1.0

Table 2. Mean RTs (ms) and accuracy (% errors) as a function of cognate status and primetype (Experiment 2, L2 pseudohomophone translation primes – L1 targets). L2 homophone translation equivalents are displayed between brackets (these words were not presented during the experiment).

Responses to L1 targets following a L2 pseudohomophone translation prime (569 ms) were significantly faster than those following a graphemic control prime (590 ms). This priming effect did not interact with cognate status (F_1 < 1, $F_2(1, 54) = 2.20$, $MSE = 9511$, $p > .14$). Indeed, Table 2 shows that the priming effect was exactly 21 ms for both cognates and non-cognates. Planned comparisons showed that the priming effect was significant both in the cognate and non-cognate conditions, respectively $F_1(1, 21) = 6.10$, $MSE = 801$, $p < .01$, $F_2(1, 54) = 532.26$, $MSE = 9511$, $p < .001$ and $F_1(1, 21) = 3.45$, $MSE = 1448$, $p < .04$, $F_2(1, 54) = 439.50$, $MSE = 9511$, $p < .001$ (one-tailed tests). Similar to Experiment 1, there were no effects of cognate status or primetype on accuracy (all Fs < 1).

Comparison Experiment 1 and 2. In order to compare the strength of the priming effect for L1 and L2 primes (respectively L2 and L1 targets), we also analyzed the data from Experiment 1 and 2 as one design. Again, an ANOVA was performed with Cognate Status and Primetype as repeated measures factors. In addition, Language of the prime/target (L1-L2 vs. L2-L1) was included as a between-subjects variable. The dependent variable was the mean RT across trials. As expected, responses were significantly slower to L2 targets (673 ms) than to L1 targets (580 ms), $F_1(1, 42) = 9.77$, $MSE = 38857$, $p < .01$, $F_2(1, 108) = 53.45$, $MSE = 10094$, $p < .001$. More importantly, the significant pseudohomophone translation priming effect did not interact significantly with prime/target language, $F_1(1, 42) = 1.29$, $MSE = 1832$, $p > .26$, $F_2(1, 108) = 1.16$, $MSE = 5769$, $p > .28$, although inspection of Table 1 and 2 shows that it tended to be larger for L1 primes (36 ms) than for L2 primes (21 ms).

DISCUSSION

In line with expectations, we found a significant backward (L2-L1) pseudohomophone translation priming effect: L1 target words were recognized faster if they were preceded by a L2 pseudohomophone of their L2 translation equivalents. In line with Van Wijnendaele and Brysbaert (2002), this shows that the L2 pseudohomophone primes were phonologically coded through L2 GPC rules, even though the task only involved L1 target words. Moreover, these phonological representations were activated strongly enough to pre-activate the underlying semantic representations and corresponding L1 translation equivalents. Again, this effect did not interact with the degree of form overlap (cognate status) between the translation equivalents. Also, the effect of L1 pseudohomophone translation primes on L2 targets (36 ms) was not significantly larger than the effect of L2 pseudohomophones on L1 targets (21 ms), although raw means show a tendency towards such an interaction. For further theoretical implications of these findings, we refer to the *General Discussion.*

EXPERIMENT 3

METHOD

Participants. The participants were 20 Dutch-English bilinguals. Mean age was 20.09 years (SD = 3.65). None of them participated in one of the previous experiments. They belonged to the same population and had a similar L2 history as the participants in Experiment 1 and 2.

Stimulus Materials. The stimulus list was similar to Experiment 1, except that the pseudohomophone primes were not homophone to the translation equivalent of the target, but to a related word of the target (e.g. *pous* [paus-pope] – CHURCH). Also, because it was hard to find stimuli satisfying the different constraints outlined below, cognate status was not included as a stimulus variable. The stimuli consisted of 34 L2 (English) word targets and 34 L2 nonword targets. All word targets were matched with two types of L1 (Dutch) nonword primes (see Appendix C). The first type of primes were pseudohomophone associative primes, i.e. L1 nonwords which have the same pronunciation (e.g. *pous*) as the L1 translation equivalent (e.g. *paus*) of a word (e.g. pope) that is related to the L2 target (e.g. CHURCH). All associated word pairs were drawn from the free association norms database from Nelson and McEvoy (1998), which lists the (directional) strength of the associations between more than 4000 words, measured as the chance that somebody produces a certain word as the first response to a given word (and vice versa) when asked to give the first word that comes to mind. We selected those associated word pairs from the database from which the associative strength was as strong as possible, provided it was still possible to find a cross-lingual Dutch pseudohomophone of one of the words. Mean associative strength between the selected associates was .353 (with a maximum of .819 for the *day – night* pair). Similar, to the previous experiments, the second type of primes were L1 nonword graphemic control primes, matched with the pseudohomophone associative primes following the criteria and procedure described in the *Method* section of Experiment 1. The resulting set of pseudohomophone and pronounceable (in L1) control primes were matched for summated bigram frequency (respectively M = 22389 and M = 21520, $F < 1$), neighbourhood size (M = 5.36 and M = 4.53, $F < 1$), word length (identical) and orthographic overlap with the target. Care was also taken that no control prime sounded like an existing Dutch or English word. Also, English (L2) targets did not sound like, or were homographs of existing Dutch words. Mean CELEX log frequency per million of the target words was 1.87 (SD = 0.54). Mean word target length was 4.7 letters.

Procedure. The procedure was identical to that of Experiment 1, except for the number of trials. All participants completed 34 word and 34 nonword trials; again, half of the word targets was preceded by a pseudohomophone associative prime, whereas the other half was preceded by a control prime. Also, the presence of (L1) primes in the experiment was not mentioned.

RESULTS

The proportion of false responses to L2 word and nonword targets was 10.2%. These trials were discarded from all analyses. A repeated measures ANOVA was performed with Primetype (graphemic control vs. pseudohomophone) as the only independent variable. The dependent variable was the mean RT across trials. Mean RTs and accuracy as a function of primetype are presented in the left part of Table 3.

As expected and similar to Experiment 1, the effect of primetype was significant both in the analysis by participants and by items, respectively $F_1(1, 19) = 4.72$, $MSE = 2328$, $p < .05$ and $F_2(1, 33) = 4.24$, $MSE = 4121$, $p < .05$. Responses to L2 targets following a pseudohomophone of a semantically related L1 word (659 ms) were significantly faster than responses to L2 targets following a graphemic control prime (692 ms). As for accuracy, there tended to be more errors in the pseudohomophone

	L1 Prime – L2 Target			L2 Prime – L1 Target		
	Example	RT	% errors	Example	RT	% errors
Orthographic Control Prime	*zeun* – CHURCH	692	5.5	*sned* – BEEN	596	2.8
Pseudohomophone Associative Prime	*pous* [paus]– CHURCH	659	9.6	*knea* [knee] – BEEN	576	3.1
Net Priming Effect		33	-4.1		20	-0.3

Table 3. Mean RTs (ms) and accuracy (% errors) as a function of prime/target language and primetype (Experiment 3, L1 pseudohomophone associative primes - L2 targets; Experiment 4, L2 pseudohomophone associative primes – L1 targets). Homophone associatively related words are displayed between brackets (these words were not presented during the experiment).

condition than in the control condition, in contrast with our expectations and with the pattern observed in the RTs. However, this trend was not significant, $F_1(1, 19) = 3.59$, $MSE = 46.2$, $p > .07$ and $F_2(1, 33) = 3.74$, $MSE = 47.6$, $p > .06$.

Similar to the analysis in which the pseudohomophone translation effect was compared for Experiment 1 (L2 targets) and Experiment 2 (L1 targets), we also compared the magnitude of the primetype effect for pseudohomophone translation primes (Experiment 1) and pseudohomophone associative primes. The effect was somewhat (3 ms) larger for translation primes, but this difference was not significant (both *F*s < 1).

DISCUSSION

Similar to Experiment 1, we found a significant forward pseudohomophone associative priming effect. L2 target words were faster recognized if they were preceded by a L1 pseudohomophone of a word related to their L1 translation equivalents. In line with Brysbaert et al. (1999), this strongly suggests that the L1 pseudohomophone primes were phonologically coded through L1 GPC rules, even though the task only involved L2 target words. Moreover, these phonological representations were activated strongly enough to pre-activate their underlying semantic representations, which in turn activated related semantic representations and their corresponding L2 lexical entries. Also, the pseudohomophone associative priming effect found in this experiment was not significantly smaller than the pseudohomophone translation priming effect (Experiment 1). For further theoretical implications of these findings, we refer to the *General Discussion.*

EXPERIMENT 4

METHOD

Participants. The participants were 20 Dutch-English bilinguals. Mean age was 22.2 years (SD = 4.86). None of them participated in one of the previous experiments. They belonged to the same population and had a similar L2 history as the participants in Experiments 1, 2 and 3.

Stimulus Materials. The composition of the stimulus list was identical to Experiment 3, but language of the primes and targets was switched. Also, there were now 36 L1 (Dutch) word targets and 36 L2 nonword targets. Again, all word targets were matched with two types of L2 (English) nonword primes (see Appendix D). The first type of primes were pseudohomophone associative primes, i.e. L2 nonwords which have the same pronunciation (e.g. *mowse*) as the L2 translation equivalent (e.g. mouse) of a word (e.g. muis) that is related to the L1 target (e.g. KAT [cat]). Just as in Experiment 3, all associated word pairs were drawn from the free association norms database of Nelson and McEvoy (1998), on the condition that it was possible to find a cross-lingual English pseudohomophone of one of the words. Like in Experiment 2, only pseudohomophones were selected which are in the ARC nonword database (Rastle et al., 2002). Mean associative strength between the selected associates was .313 (with a maximum of .828 for the *toad – frog* pair), not differing from the mean association strength of Experiment 3, $F < 1$. Similar to the previous experiments, the second type of primes were L2 nonword graphemic control primes, matched with the pseudohomophone associative primes following the criteria and procedure described in the *Method* section of Experiment 1. The resulting set of pseudohomophone and pronounceable (in L2) control primes were matched for summated bigram frequency (respectively M = 5382 and M = 5495, $F < 1$), neighbourhood size (M = 4.97 and M = 5.31, $F < 1$), word length (identical) and orthographic overlap with the target. Care was also taken that no control prime sounded like an existing Dutch or English word. Also, English (L1) targets did not sound like, or were homographs of existing Dutch words. Mean CELEX log frequency per million of the target words was 1.68 (SD = 0.53). Mean word target length was 4.7 letters.

Procedure. The procedure was identical to that of Experiment 3, except for the number of trials. All participants completed 36 word and 36 nonword trials; again, half of the word targets was preceded by a pseudohomophone associative prime, whereas the other half was preceded by a control prime. Also, the presence of (L2) primes in the experiment was not mentioned.

RESULTS

The proportion of false responses was 3.9%. These trials were discarded from all analyses. A repeated measures ANOVA was performed with Primetype (graphemic control vs. pseudohomophone) as the only independent variable. The dependent variable was the mean RT across trials. Mean RTs and accuracy as a function of primetype are presented in the right part of Table 3.

Similar to Experiment 3 (L1 primes – L2 targets), responses to L1 targets following a pseudohomophone of a semantically related L2 word (576 ms) were somewhat faster than responses to L1 targets following a graphemic control prime (596 ms). However, this 20 ms effect was not statistically reliable, $F_1(1, 19) = 2.69$, $MSE = 1508$, $p > .11$ and $F_2(1, 35) = 2.56$, MSE

=2320, $p > .11$. As can already be seen in Table 3, there was no effect of primetype on accuracy (both Fs < 1).

DISCUSSION

Whereas the pseudohomophone translation priming effect was equally strong from L2 primes to L1 targets as vice versa (Experiments 2 and 1), this was not the case for the cross-lingual pseudohomophone associative priming effect. In contrast with the strong effect of L1 primes/L2 targets obtained in Experiment 3, the effect observed in this Experiment (L2 primes/L1 targets) was not statistically reliable. Responses to L1 target words following L2 pseudohomophone associative primes were not significantly faster than those following L2 control primes (although there was a 20 ms effect). This suggests that the L2 pseudohomophones were possibly phonologically coded to some degree (given the results of Van Wijnendaele & Brysbaert, 2002), but these phonological representations were not activated strongly enough to pre-activate their underlying semantic representations and/or related semantic representations and their corresponding L1 lexical entries. For further theoretical implications of these findings, we refer to the *General Discussion.*

EXPERIMENT 5

METHOD

Participants. The participants were 22 Dutch-English bilinguals. Mean age was 21.23 years (SD = 4.44). None of them participated in one of the previous experiments. They belonged to the same population and had a similar L2 history as the participants in Experiments 1 to 4.

Stimulus Materials. The stimuli consisted of 23 L2 (English) word targets and 23 L2 nonword targets. All word targets were matched with two types of L2 (English) nonword primes (see left part of Appendix E). The first type of primes were intra-lingual homophone intermediate translation primes, i.e. L2 words (e.g. *hook*) which have the same pronunciation as the L1 translation equivalent (e.g. hoek) of the L1 target (e.g. CORNER). The second type of primes were L2 graphemic control primes, matched with the homophones analogue to the criteria and procedure described in the *Method* section of Experiment 1. The resulting set of homophones and control primes were matched for word frequency (respectively M = 1.67 and M = 1.57, $F < 1$) summated bigram frequency[4] (M = 7445 and M = 7042, $F < 1$), neighbourhood size (M = 12.83 and M = 9.91, $p > .11$), word length (identical) and orthographic overlap with the target. Care was also taken that no words were existing words in Dutch (L1). The 23 pronounceable L2 nonword targets satisfied the criteria mentioned above for the nonword target stimuli in Experiment 1. Mean log frequency per million of the L2 target words was 1.53 (SD = 0.65).

Procedure. The procedure was identical to that of Experiment 1. Each participant completed 56 trials. 12 or 11 of the 23 word targets (counterbalanced over subjects) were preceded by a homophone intermediate translation prime. The other word targets were preceded by a control prime. Again, each participant received a different prime permutation. Across participants, all targets were displayed with the two types of primes. Again, the presence of (L2) primes in the experiment was not mentioned. The

[4] Note that these measures of summated bigram frequency (and to a lesser extent of neighbourhood size) can not directly be compared with the respective measures for the L1 nonwords of Experiment 1. The latter are much higher because the Dutch CELEX contains much more records than the English CELEX, leading to a higher overall number of bigram occurrences. For a more detailed discussion of this topic, we refer to Duyck et al. (2004).

intermediate L1 translation equivalents (or any other L1 words) were not displayed during the experiment.

RESULTS

The proportion of false responses was 13.1%. Again, these trials were excluded from all analyses. Mean RTs and proportion of correct trials as a function of primetype are presented in the left part of Table 4. A repeated measures ANOVA was performed with Primetype (graphemic control vs. homophone of translation equivalent) as the only independent variable. The dependent variable was the mean RT across trials.

	L2Primes/Targets			**L1 Primes/Targets**		
	Example	RT	% errors	Example	RT	% errors
Orthographic Control Prime	*foot* – CORNER	936	9.8	*dek* – DAG	724	4.8
Homophone [Intermediate Translation] Prime	*hook* [hoek]– CORNER	880	12.3	*dij* [day] – DAG	731	3.8
Net Priming Effect		56	-2.5		-7	1.0

Table 4. Mean RTs (ms) and accuracy (% errors) as a function of target language and primetype (Experiment 5 and 6). Intermediate homophone translation equivalents are displayed between brackets (these words were not presented during the experiment).

Interestingly, this showed a large significant effect of Primetype, $F_1(1, 21) = 4.53$, $MSE = 7360$, $p < .05$. Responses to targets following an intra-lingual homophone of its translation equivalent (880 ms) were 56 ms faster than responses to control primes (936 ms). This effect however, although very large, was not significant in the item analysis, $F_2 < 1$. Finally, there was no effect of primetype on the proportion of correct responses, $F_1(1, 21) = 1.14$, $MSE = 64.5$, $p > .29$, $F_2(1, 22) = 1.76$, $MSE = 69.5$, $p > .19$.

Because the effect of primetype was very large and reliable in the analysis by participants, but by far not significant in the analysis by items, there is reason to believe that the primetype effect interacts with some item variable not accounted for in the experimental design, and was therefore only present for some of the stimuli. To further explore this hypothesis, we have repeated the analysis described above, including one additional independent variable, which we believed might have interacted with the primetype effect. Because word frequency is probably the linguistic variable with the most robust effects in the psycholinguistic literature, we decided to include the relative word frequency of the target compared to that of the primes. Whereas homophone and control primes were matched for word frequency, this was not the case for target frequency relative to prime frequency, because there are only very few stimuli which are homophone to the other's translation equivalent (in addition to the other selection criteria mentioned above). Hence, the stimulus set contained triplets (n = 11) for which the target had a lower frequency compared to the primes (e.g. THIGH; day/new), whereas the opposite was true for other triplets (n = 12, e.g. TIME; tide/tile). This variable had no effect on overall RTs, $F_1(1, 21) = 2.23$, $MSE = 20437$, $p > .15$, $F_2 < 1$. However, interestingly enough, this factor interacted with the primetype effect, $F_1(1, 21) = 4.46$, $MSE = 16376$, $p < .05$, $F_2(1, 21) = 3.77$, $MSE = 7767$, $p < .07$. The priming effect was 113 ms when the frequency of the target was lower than that of the primes, whereas the effect was -2 ms when the opposite was true. Planned comparisons showed that this first difference was significant both in the analysis by participants and by items ($F_1(1, 21) = 4.80$ $MSE = 29338.2$, $p < .05$, $F_2(1, 21) = 4.40$, $MSE = 7766.72$, $p < .05$), whereas the second was not (both Fs < 1). Note that this interaction effect of primetype and relative prime/target frequency is not a confounded effect of target frequency (i.e. a primetype effect for low frequent, but not for high frequent targets). An analysis with this factor instead of the relative frequency factor yielded no significant results.

DISCUSSION

As mentioned earlier, the present experiment was set up to find evidence for (a) pre-lexical phonological coding of L2 primes during L2 processing and (b) for language-independent semantic activation of phonological representations (i.e. does the pre-lexically assembled phonological representation of *hook*, /huk/, activates both its L1 and L2 meaning?). The significant 113 ms priming effect for intra-lingual translation equivalent homophones is supportive of both statements. Hence, it is very likely that the prime *hook* was pre-lexically phonologically coded. This phonological representation (/huk/) activated both its L1 [corner] and L2 [hook] meaning. The former led to pre-activation of the associated L2 lexical label for that meaning, which caused the prime effect observed. An important qualification to this line of reasoning concerns the fact this strong effect was only found when the L2 prime was more frequent than the L2 target. It is unclear whether this factor mainly influenced the degree of pre-lexical phonological coding of the L2 prime, or the strength of the phonology (/huk/) to meaning [corner] mapping[5]. For further theoretical implications of these findings, we refer to the *General Discussion.*

---

[5] It could be argued that the strength of the mapping between /huk/ and the meaning [corner] is (at least partially) correlated with the frequency by which /huk/ occurs in a language to indicate the meaning [corner] (i.e. the mapping will be stronger for high frequent than for low frequent phonological representations). As /huk/ is used in Dutch to indicate the meaning [corner], this would coincide with the Dutch spoken word frequency of /huk/. Because spoken word frequencies are not available for Dutch, written word frequencies of the corresponding words (*hoek* in this case) probably are good approximate measures (for the English CELEX, we calculated the correlation between available spoken and written word frequencies; this was 0.87, $p < .001$). Now, an analysis similar to the relative prime/target frequency analysis, including these frequencies instead, indicated that this variable did not influence the primetype effect at all and is therefore probably not responsible for the primetype x relative frequency interaction. This suggests that this interaction effect observed is probably due to weaker activation of phonological representations in low frequent L2 primes. But, of course, this line of reasoning is by no means a definite explanation for the issue at hand.

EXPERIMENT 6

METHOD

Participants. The participants were 24 Dutch-English bilinguals. Mean age was 23.23 years (SD = 6.44). None of them participated in one of the previous experiments. They belonged to the same population and had a similar L2 history as the participants in Experiments 1 to 5.

Stimulus Materials. The composition of the stimulus list was identical to that of Experiment 5, but language was switched. The stimuli consisted of 23 L1 (Dutch) word targets and 23 L1 nonword targets. All word targets were matched with two types of L1 nonword primes (see right part of Appendix E). The first type of primes were intra-lingual homophone intermediate translation primes, i.e. L1 words (e.g. *bijl* [axe]) which have the same pronunciation as the L2 translation equivalent (e.g. bail) of the L1 target (e.g. BORG). The second type of primes were L1 graphemic control primes, matched with the homophones analogue to the criteria and procedure described in the *Method* section of Experiment 1. The resulting set of homophones and control primes were matched for word frequency (respectively M = 1.23 and M = 1.26, F < 1) summated bigram frequency[4] (M = 21900 and M = 21288, F < 1), neighbourhood size (M = 12.35 and M = 10.70, F < 1), word length (identical) and orthographic overlap with the target. Care was also taken that no words were also existing words in English (L2). The 23 pronounceable L1 nonword targets satisfied the criteria mentioned above for the nonword target stimuli in Experiment 1. Mean log frequency per million of the L1 target words was 1.40 (SD = 0.87).

Procedure. The procedure was identical to that of Experiment 1. Each participant completed 56 trials. 12 or 11 of the 23 word targets (counterbalanced over subjects) were preceded by a homophone intermediate translation prime. The other word targets were preceded by a control prime. Again, each participant received a different prime permutation. Across participants, all targets were displayed with the two types of primes. Again, the presence of (L2) primes in the experiment was not mentioned. The intermediate L1 translation equivalents (or any other L1 words) were not displayed during the experiment.

RESULTS

The proportion of false responses was 7.1%. Again, these trials were excluded from all analyses. Mean RTs and proportion of correct trials as a function of primetype are presented in the right part of Table 4. A repeated measures ANOVA was performed with Primetype (graphemic control vs. homophone of translation equivalent) as the only independent variable. The dependent variable was the mean RT across trials. In contrast with the previous experiment (L2 targets), the effect of primetype was far from significant (both Fs < 1). Responses to L1 targets following an intra-lingual homophone of its translation equivalent were even slightly slower than responses to control primes, respectively 731 ms and 724 ms. Just as in the previous experiments, there was also no effect of primetype on the proportion of correct responses (both Fs < 1).

In accordance with Experiment 5, we also ran the same analysis including the relative frequency of the target compared to the primes. In contrast with the previous experiment, this factor had no effect ($F < 1$ for all main and interaction effects). The primetype effect was -4 ms for targets with higher word frequency than their primes, and -8 ms for targets having lower frequency.

DISCUSSION

In contrast with Experiment 5, no effects were found of intra-lingual translation equivalent homophones on the processing of L1 targets. Given the large body of evidence supporting the claim of pre-lexical phonological coding of L1 words (see the introduction), it is unlikely that the absence of this effect is due to the fact that the L1 primes might not have activated their phonological representations. Instead, it is more probable that this was caused by the fact that the mapping from an ambiguous phonological code (e.g. /dei/) on its L2 meaning [day] is weaker than the mapping from phonology on L1 meaning [thigh]. For further theoretical implications of these findings, we refer to the *General Discussion*.

GENERAL DISCUSSION

During the last decade, a strong phonological model of monolingual word recognition has gained importance. In this model, it is assumed that words are coded phonologically before lexical access takes place. Recently, Brysbaert et al. (1999) have shown that these processes generalize across languages in multilinguals to a certain extent. They found that L2 words (e.g. *pour* [for]) are recognized faster if they are preceded by a masked L1 pseudohomophone (e.g. *poer*) prime. Later, Van Wijnendaele and Brysbaert have replicated this effect with L1 targets and L2 primes (but see Jared & Kroll, 2001). These findings offer strong evidence that L1 GPC rules are processed during L2 processing and vice versa. The goal of this paper was to further investigate phonological coding in bilinguals, by extending the monolingual pseudohomophone priming effect (e.g. Grainger & Ferrand, 1996; Lukatela & Turvey, 1994b; Lukatela et al., 1998; Perfetti & Bell, 1991) and the pseudohomophone associative priming effect (e.g. Drieghe & Brysbaert, 2002; Lesch & Pollatsek, 1993; Lukatela & Turvey, 1994a) to mixed-language stimuli.

Experiments 1 and 2 yielded the strongest evidence for our claim. First, we showed that the recognition of L2 words (e.g. BACK) is significantly facilitated by a L1 masked prime which is a pseudohomophone of its translation equivalent (e.g. *ruch* [rug]). Secondly, we showed that the same applies for L1 targets (e.g. TOUW) and L2 pseudohomophone translation primes (e.g. *roap* [rope]; control prime *joll*). Because the primes were only homophone to the translation equivalent according to GPC rules of the non-target language, this strongly suggests that L1 GPC rules are active during L2 word recognition, and vice versa. This is in line with the respective studies of Brysbaert et al. (1999) and Van Wijnendaele and Brysbaert (2002). The prime effect was equally large from L2 to L1 as from L1 to L2, and the effects did not interact with form overlap (cognate status) between the two respective translation equivalents. Overall, this pseudohomophone translation effect is compatible with findings of Tan and Perfetti (1997; but see Zhou & Marslen-Wilson, 1999), who demonstrated that a Chinese target word can be primed with a homophone of a target synonym. If one starts from a strong non-selective view on lexical access in bilinguals (see earlier, Dijkstra & Van Heuven, 2002), there is not much difference between an 'intra-lingual' synonym and a 'cross-language' translation equivalent, in that they are both different lexical labels representing the same meaning. Also, in contrast with the results of Jared and Kroll (2001), these effects suggest that L2 phonology can influence L1 processing, even if the participants did not just read words in L2.

Note that the masked priming paradigm we used excludes the possibility that these effects were due to strategic (target 'guessing') factors. Hutchinson, Neely and Johnson (2001; see also Neely, 1977; Neely, 1991) for example showed that such factors do not influence semantic priming when the SOA between primes target is 167 ms. In our study, this was only 114 ms (a 57 ms prime followed by a 57 ms postmask). The fact that these effects were obtained with a lexical decision task, which does not explicitly require phonology (unlike for example a naming task), adds further strength to the automaticity of these processes.

In experiment 3, we replicated the above effect with L2 targets (e.g. CHURCH) and L1 pseudohomophone associative primes (e.g. *pous* [paus – pope]) (instead of pseudohomophone translation primes). Again, this strongly suggests L1 phonological coding during L2 processing. Also, this shows that the overlap between the concepts that the pseudohomophone and the target represent needs not to be complete (as is the case for translation equivalents) in order for the pseudohomophone priming effect to arise. This shows that the activation in the L1 phonological representations is quite large, certainly large enough to spread to related concepts (for a more detailed account of this effect, see further). In contrast with the symmetry between Experiment 1 and 2 however, this pseudohomophone associative priming effect was not significant with L1 targets (e.g. BEEN [leg]) and L2 primes (e.g. *knea*), although we observed a 20 ms effect in the expected direction. Possibly, the L2 pseudohomophones were phonologically coded (given the results of Van Wijnendaele & Brysbaert, 2002, and Experiment 2, this study) to a certain extent, but these phonological representations were not activated strongly enough to pre-activate their underlying semantic representations and/or related semantic representations and their corresponding L1 lexical entries. On the other hand, the results of Jared and Kroll (2001), who found that L2 phonology only influences L1 naming when L2 GPC rules have recently been active, suggest that the weakness of the effect here may be due to the low rest activation in the L2 phonological processing system (participants were not aware of the bilingual nature of any of our experiments).

Finally, in the last two experiments, we tried to find indications of cross-lingual phonological influences in a monolingual stimulus context. In Experiment 5, we found that the recognition of L2 words (e.g. CORNER) is facilitated by L2 homophones (e.g. *hook*) of their L1 translation equivalents (e.g. hoek). First, this shows that words are also pre-lexically phonologically coded when reading in L2 (see also Brysbaert et al., 1999, Experiment 1). Second, this shows that ambiguous L2 phonological representations (interlingual homophones, e.g. /huk/) quickly activate all underlying semantic representations, even if they correspond to two different languages and are not related (e.g. [hook]-[corner]). In this case for instance, the phonological representation /huk/ activated its L1 meaning [corner], even though the experiment only contained L2 stimuli. However, it is important to note that the prime was only able to influence target recognition if it was more frequent than the target (although the effect was significant in the analysis by participants across all stimuli). It is unclear whether this relative frequency mainly influenced the degree of pre-lexical phonological coding of the L2 prime, or the strength of the phonology to meaning [corner] mapping (see *Footnote 2*). In Experiment 6, we observed no significant effects: L1 targets (e.g. DAG) were not processed faster if they were preceded by intra-lingual homophones (e.g. *dij*) of their L2 translations (day). Given the large body of evidence for pre-lexical phonological coding in L1 word recognition (see the *Introduction*), the absence of an effect here probably is due to the fact that the mapping from an ambiguous phonological code (e.g. /dei/) on its L2 meaning [day] is much weaker than the mapping from phonology on the L1 meaning (e.g. *dij*) [thigh].

As noted in the very beginning of this paper, research on language-selective functioning of the bilingual language processing system has mainly focused on lexical representations (e.g. Dijkstra & Van Heuven, 2002). As a consequence, there is no model of bilingual phonological processing at present. It is clear from the present and previous research (Brysbaert et al., 1999; Van Wijnendaele & Brysbaert, 2002; Jared & Kroll, 2001), that any future model will have to be structurally language non-selective with regard to the activation of phonological representations, much in the way that the Bilingual Interactive Activation Plus (BIA+) model (Dijkstra & Van Heuven, 2002) is non-selective for lexical access. This model is an extension of the Interactive Activation model for monolingual word recognition (e.g. McClelland & Rumelhart, 1981), containing language, word, letter and feature nodes. In the model, all L2 and L1 words are represented into a unitary word-level system. Lexical access during word recognition is initially non-selective, as word activation is affected by competing items from both languages. Because the model (unlike the earlier BIA model) does not contain any top-down connections, effects of language context and stimulus list composition (e.g. Dijkstra et al., 2000) are dealt with at the task schema level, which only receives input from the (fundamentally language non-selective) word identification system.

More importantly for the present study, the sketch of the recent BIA+ model also contains semantic and phonological representations, although these have not been implemented yet and are basically still black-boxes. It will be very interesting to see whether this model will be able to cope with the results of this and previously mentioned studies, if the phonological subsystem is also conceived as fundamentally non-selective and highly interacting with semantic and lexical representations. At present, probably the most important assumption that the authors have made with respect to this subsystem is the temporal delay assumption. This states that L2 phonological (and semantic) representations are delayed in activation relative to L1 codes. The present study suggests that this assumption may be too strong (see also Brysbaert, Van Wijnendaele, & Duyck, 2002) and that relative speed of phonological activation may be less language-dependent. For instance, the cross-lingual pseudohomophone translation priming effect was not significantly stronger from L1 to L2 (Experiment 1) than from L2 to L1 (Experiment 2). Also, Experiment 5 showed a priming effect of L2 primes (homophones of the L2 target's translation equivalent), whereas no such effect was found in Experiment 6 (L1 primes), Also, the homophonic priming effect of Van Wijnendaele and Brysbaert (2002, see earlier) was equally large for L2 as for L1. These observations all suggest that L2 phonological processing is not always necessarily running behind on L1 phonological activation. However, the findings of Jared and Kroll (2001) suggest that the speed of L2 phonological processing and its impact on L1 processing may be very sensitive to recent use of L2 (such as naming a block of L2 filler words) due to a lower rest activation level relative to L1. On the other hand, we did find evidence for the temporal delay assumption with respect to the connections from phonological to semantic representations. Experiment 5 showed that ambiguous phonological representations (from interlingual homophones) always activate their L1 meaning, even when performing a L2 task, whereas there was no sign of those phonological representations activating their L2 meaning in a L1 task (Experiment 6), although the pre-lexical phonological coding of the L1 primes (Experiment 5) was probably larger than for L2 primes (Experiment 6). Another interesting issue will be whether BIA+ will be able to account for cross-linguistic differences between studies. For instance, evidence for L2 phonological coding during L1 processing is less convincing for bilinguals who's two languages have different alphabets (see earlier, Gollan et al., 1997; Kim & Davis, 2003; Zhou & Marslen-Wilson, 1999; but see Tan & Perfetti, 1997). This is consistent with the idea that in more form-related languages (such as Dutch and English), transfer of L1 knowledge (such as GPC rules) during L2 acquisition is easier because those languages contain many letter-sound combinations which are very similar (for a more detailed discussion of this issue, see Brysbaert et al., 2002).

Throughout this paper, we have assumed specified semantic involvement in both the cross-lingual pseudohomophone associative and translation effects, following the accounts of the respective monolingual effects (e.g. Frost, 1998; e.g. Lesch & Pollatsek, 1993; Lukatela & Turvey, 1994a). For instance, the forward pseudohomophone translation effect was accounted for as follows: first, the L1 nonword prime accesses its phonological code. This activates the corresponding semantic representation which in turn pre-activates the corresponding L2 lexical entry for that meaning (causing the priming effect). The same line of reasoning applies for the backward pseudohomophone translation priming effect (L2 primes/L1 targets). For pseudohomophone associative priming (e.g. Experiment 3), one additional step is required: after semantic access, activation is spread to related concepts which share semantic features. Then, again corresponding lexical entries of those related concepts are pre-activated. Although this account of the observed effects seems very plausible, it is important to point out one alternative explanation. It is possible that the pre-lexically assembled phonological code does not activate semantic representations, but rather the lexical entry which is associated with that phonological representation. Then, activation can be spread to translation equivalents (Experiment 1, 2) or related words through strong lexical links, without semantic involvement (see for example the strong lexical links between L2 translation equivalents and L1 words in the model of bilingual language organization of Kroll & Stewart, 1994). Although our data do not allow excluding this hypothesis with absolute certainty, we are inclined to situate the locus of the priming effects within the semantic system (like Lesch & Pollatsek, 1993; Lukatela & Turvey, 1994a) for a number of reasons. First, Lukatela and Turvey (1994a) showed that pseudohomophone associative primes (e.g. *tode*) are as effective as word primes (e.g. *toad*). This should not be the case if the associative priming effect occurs through lexical representations, because activation in the lexical entry should be larger if the lexical form is actually displayed than when it is only activated through phonological code. Similarly, in the homophone associative priming study of Lesch and Pollatsek (1993, see earlier), homophones of associated words (e.g. *beach*) were equally effective primes for targets (e.g. *nut*) as the associated words (e.g. *beech*) their selves. Secondly, Van Orden (1987) and Van Orden, Johnston and Hale (1988) showed that participants made more false positive errors to homophones and pseudohomophones (e.g. *rows* and *sute*) of category exemplars (e.g. rose and suit) than to control words in a semantic categorization task (e.g. flowers and clothes). This also suggests that phonological codes can rapidly access meaning. Third, the pseudohomophone translation priming effect found in Experiment 1 and 2 did not interact with the form overlap (cognate status) between the two involved translation equivalents. If these two words pre-activate each other through lexical links, one would expect a larger effect for (near) cognates. Fourth, the hypothesis that phonology accesses orthography before meaning is not only counterintuitive, it is also not compatible with the speech primacy axiom, according to which the primary association formed during language acquisition is the connection between spoken words and meaning. In this view, written language is a secondary system, appended onto the already existing system (Frost, 1998, pp. 74). Fifth, Lucas (2000) showed in a meta-analysis of 26 studies that semantic priming generally has an effect independent of association, which is nevertheless able to add an “associative boost” (Lucas, 2000, pp. 618) to a semantic relation. Finally, if these effects are indeed semantically mediated, this suggests that mappings from L2 lexical representations onto meaning may be stronger than previously thought (e.g. in the model of Kroll & Stewart, 1994), which is in line with more recent research on this issue (for a detailed discussion, see Duyck & Brysbaert, 2004; Duyck & Brysbaert, 2002; Francis et al., 2003).

In conclusion, the present study provided evidence against a strong language selective view on phonological coding in bilinguals, in line with previous research from Brysbaert et al. (1999), Van Wijnendale and Brysbaert (2002) and Jared and Kroll (2001). Using a masked priming paradigm, we showed that L2 words can be primed with L1 pseudohomophones of their translation equivalents, and vice versa. Also, we extended the pseudohomophone associative priming effect (Lukatela & Turvey, 1994a; Drieghe & Brysbaert, 2002) to L2 targets and L1 pseudohomophone associative primes. Finally, we found strong indications that interlingual homophones always activate their L1 meaning, even when performing a task in L2.

REFERENCES

Baayen, R., Piepenbrock, R., & Van Rijn, H. (1993). *The CELEX lexical database*. [CD-ROM]. Philadelphia: University of Pennsylvania: Linguistic Data Consortium.

Berent, I., & Perfetti, C. A. (1995). A Rose Is a Reez - the 2-Cycles Model of Phonology Assembly in Reading English. *Psychological Review, 102*(1) , 146-184.

Bovens, N., & Brysbaert, M. (1990). Ibm-Pc/Xt/at and Ps/2 Turbo Pascal Timing With Extended Resolution. *Behavior Research Methods Instruments & Computers, 22*(3), 332-334.

Brysbaert, M., Van Dyck, G., & Van De Poel, M. (1999). Visual Word Recognition in Bilinguals: Evidence From Masked Phonological Priming. *Journal of Experimental Psychology: Human Perception and Performance, 25*(1), 137-148.

Brysbaert, M., Van Wijnendaele, I., & Duyck, W. (2002). On the temporal delay assumption and the impact of non-linguistic context effects. *Bilingualism: Language and Cognition, 5*(3), 199-201.

Brysbaert, M., Vitu, F., & Schroyens, W. (1996). The Right Visual Field Advantage and the Optimal Viewing Position Effect: on the Relation Between Foveal and Parafoveal Word Recognition. *Neuropsychology, 10*(3), 385-395.

Carreiras, M., & Perea, M. (2002). Masked Priming Effects With Syllabic Neighbors in a Lexical Decision Task. *Journal of Experimental Psychology: Human Perception and Performance, 28*(5), 1228-1242.

Coltheart, M. (1978). Lexical Access In Simple Reading Tasks. In Underwood, G. (Ed.), *Strategies Of Information Processing.* (pp. 151-216). London: Academic Press.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). Drc: a Dual Route Cascaded Model of Visual Word Recognition and Reading Aloud. *Psychological Review, 108*(1), 204-256.

Dijkstra, A., & Van Heuven, W. (2002). The Architecture of the Bilingual Word Recognition System: From Identification to Decision. *Bilingualism: Language and Cognition, 5*(3), 175-197.

Dijkstra, T., Grainger, J., & Van Heuven, W. J. B. (1999). Recognition of Cognates and Interlingual Homographs: the Neglected Role of Phonology. *Journal of Memory and Language, 41*(4), 496-518.

Dijkstra, T., Timmermans, M., & Schriefers, H. (2000). On Being Blinded by Your Other Language: Effects of Task Demands on Interlingual Homograph Recognition. *Journal of Memory and Language, 42*(4), 445-464.

Drieghe, D., & Brysbaert, M. (2002). Strategic Effects in Associative Priming With Words, Homophones, and Pseudohomophones. *Journal of Experimental Psychology: Learning Memory and Cognition, 28*(5), 951-961.

Duyck, W., & Brysbaert, M. (2002). What Number Translation Studies Can Teach Us About the Lexico-Semantic Organisation in Bilinguals. *Psychologica Belgica, 42*(3), 151-175.

Duyck W., & Brysbaert M. (2004). *Forward and Backward Number Translation Requires Conceptual Mediation Both in Balanced and Unbalanced Bilinguals*. Manuscript submitted for publication.

Duyck W., Desmet T., Verbeke L., & Brysbaert M. (2004). *WordGen: A Tool for Word Selection and Non-Word Generation in Dutch, German, English, and French.* Manuscript submitted for publication.

Duyck, W., Diependaele, K., Drieghe, D., & Brysbaert, M. (2004). The Size of the Cross-Lingual Masked Phonological Priming Effect Does not Depend on Second Language Proficiency. *Experimental Psychology, 51*(2), 1-9.

Farrar, W. T., Van Orden, G. C., & Hamouz, V. (2001). When Sofa Primes Touch: Interdependence of Spelling, Sound, and Meaning in "Semantically Mediated" Phonological Priming. *Memory & Cognition, 29*(3), 530-539.

Ferrand, L., & Grainger, J. (1992). Phonology and Orthography in Visual Word Recognition - Evidence From Masked Nonword Priming. *Quarterly Journal of Experimental Psychology Section, 45A*(3), 353-372.

Ferrand, L., & Grainger, J. (1993). The Time-Course of Orthographic and Phonological Code Activation in the Early Phases of Visual Word Recognition. *Bulletin of the Psychonomic Society, 31*(2), 119-122.

Francis, W. S., Augustini, B. K., & Saenz, S. P. (2003). Repetition Priming in Picture Naming and Translation Depends on Shared Processes and Their Difficulty: Evidence From Spanish-English Bilinguals. *Journal of Experimental Psychology: Learning Memory and Cognition, 29*(6), 1283-1297.

Frost, R. (1998). Toward a Strong Phonological Theory of Visual Word Recognition: True Issues and False Trails. *Psychological Bulletin, 123*(1), 71-99.

Frost, R., Ahissar, M., Gotesman, R., & Tayeb, S. (2003). Are Phonological Effects Fragile? The Effect of Luminance and Exposure Duration on Form Priming and Phonological Priming. *Journal of Memory and Language, 48*(2), 346-378.

Gollan, T. H., Forster, K. I., & Frost, R. (1997). Translation Priming With Different Scripts: Masked Priming With Cognates and Noncognates in Hebrew-English Bilinguals. *Journal of Experimental Psychology: Learning Memory and Cognition, 23*(5), 1122-1139.

Grainger, J., & Ferrand, L. (1996). Masked Orthographic and Phonological Priming in Visual Word Recognition and Naming: Cross-Task Comparisons. *Journal of Memory and Language, 35*(5), 623-647.

Humphreys, G. W., Evett, L. J., & Taylor, D. E. (1982). Automatic Phonological Priming in Visual Word Recognition. *Memory & Cognition, 10*(6), 576-590.

Hutchison, K. A., Neely, J. H., & Johnson, J. D. (2001). With Great Expectations, Can Two "Wrongs" Prime a "Right"? *Journal of Experimental Psychology: Learning Memory and Cognition, 27*(6), 1451-1463.

Jared, D., & Szucs, C. (2002). Phonological activation in bilinguals: Evidence from interlingual homograph naming. *Bilingualism: Language and Cognition, 5*(3), 225-239.

Jared, D., & Kroll, J. F. (2001). Do Bilinguals Activate Phonological Representations in One or Both of Their Languages When Naming Words? *Journal of Memory and Language, 44*(1), 2-31.

Kim, J., & Davis, C. (2002). Using Korean to Investigate Phonological Priming Effects Without the Influence of Orthography. *Language and Cognitive Processes, 17*(6), 569-591.

Kim, J., & Davis, C. (2003). Task Effects in Masked Cross-Script Translation and Phonological Priming. *Journal of Memory and Language, 49*(4), 484-499.

Kroll, J. F., & Stewart, E. (1994). Category Interference in Translation and Picture Naming - Evidence for Asymmetric Connections Between Bilingual Memory Representations. *Journal of Memory and Language, 33*(2), 149-174.

Lesch, M. F., & Pollatsek, A. (1993). Automatic Access of Semantic Information by Phonological Codes in Visual Word Recognition. *Journal of Experimental Psychology: Learning Memory and Cognition, 19*(2), 285-294.

Lucas, M. (2000). Semantic Priming Without Association: a Meta-Analytic Review. *Psychonomic Bulletin & Review, 7*(4), 618-630.

Lukatela, G., Carello, C., Savic, M., Urosevic, Z., & Turvey, M. T. (1999). When Nonwords Activate Semantics Better Than Words. *Cognition, 69*(3), B31-B40

Lukatela, G., Frost, S. J., & Turvey, M. T. (1998). Phonological Priming by Masked Nonword Primes in the Lexical Decision Task. *Journal of Memory and Language, 39*(4), 666-683.

Lukatela, G., Savic, M., Urosevic, Z., & Turvey, M. T. (1997). Phonological Ambiguity Impairs Identity Priming in Naming and Lexical Decision. *Journal of Memory and Language, 36*(3), 360-381.

Lukatela, G., & Turvey, M. T. (1994a). Visual Lexical Access Is Initially Phonological .1. Evidence From Associative Priming by Words, Homophones, and Pseudohomophones. *Journal of Experimental Psychology: General, 123*(2), 107-128.

Lukatela, G., & Turvey, M. T. (1994b). Visual Lexical Access Is Initially Phonological .2. Evidence From Phonological Priming by Homophones and Pseudohomophones. *Journal of Experimental Psychology: General, 123*(4), 331-353.

McClelland, J. L., & Rumelhart, D. E. (1981). An Interactive Activation Model of Context Effects in Letter Perception .1. An Account of Basic Findings. *Psychological Review, 88*(5), 375-407.

Nas, G. (1983). Visual Word Recognition in Bilinguals - Evidence for a Cooperation Between Visual and Sound Based Codes During Access to a Common Lexical Store. *Journal of Verbal Learning and Verbal Behavior, 22*(5), 526-534.

Neely, J. (1991). Semantic Priming Effects in Visual Word Recognition: A Selective Review of Current Findings and Theories. In Besner, D. and Humphreys, G. W. (Eds.), *Basic Processes in Reading: Visual Word Recognition.* (pp. 264-336). Hillsdale, NJ: Erlbaum.

Neely, J. H. (1977). Semantic Priming and Retrieval From Lexical Memory - Roles of Inhibition-Less Spreading Activation and Limited-Capacity Attention. *Journal of Experimental Psychology: General, 106*(3), 226-254.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. http://www.usf.edu/FreeAssociation/.

Perea, M., & Lupker, S. J. (2003). Does Jugde Activate Court? Transposed-Letter Similarity Effects in Masked Associative Priming. *Memory & Cognition, 31*(6), 829-841.

Perfetti, C. A., & Bell, L. (1991). Phonemic Activation During the 1st 40 Ms of Word Identification - Evidence From Backward-Masking and Priming. *Journal of Memory and Language, 30*(4), 473-485.

Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 Nonwords: the Arc Nonword Database. *Quarterly Journal of Experimental Psychology, 55A*(4), 1339-1362.

Taft, M., & Van Graan, F. (1998). Lack of Phonological Mediation in a Semantic Categorization Task. *Journal of Memory and Language, 38*(2), 203-224.

Tan, L. H., & Perfetti, C. A. (1997). Visual Chinese Character Recognition: Does Phonological Information Mediate Access to Meaning? *Journal of Memory and Language, 37*(1), 41-57.

Van Hell, J. G., & Dijkstra, T. (2002). Foreign Language Knowledge Can Influence Native Language Performance in Exclusively Native Contexts. *Psychonomic Bulletin & Review, 9*(4), 780-789.

Van Orden, G. C. (1987). A Rows Is a Rose - Spelling, Sound, and Reading. *Memory & Cognition, 15*(3), 181-198.

Van Orden, G. C., Johnston, J. C., & Hale, B. L. (1988). Word Identification in Reading Proceeds From Spelling to Sound to Meaning. *Journal of Experimental Psychology: Learning Memory and Cognition, 14*(3), 371-386.

Van Wijnendaele, I., & Brysbaert, M. (2002). Visual Word Recognition in Bilinguals: Phonological Priming from the Second to the First Language. *Journal of Experimental Psychology: Human Perception and Performance, 3*, 619-627.

Zhou, X. L., & Marslen-Wilson, W. (1999). Phonology, Orthography, and Semantic Activation in Reading Chinese. *Journal of Memory and Language, 41*(4), 579-606.

APPENDIX A

Stimuli Experiment 1: L2 Targets with their Respective L1 Pseudohomophone Translation Primes and Graphemic Control Primes

Non-Cognates				Cognates			
Target	Translation Equivalent*	Pseudohomophone Translation Prime	Graphemic Control Prime	Target	Translation Equivalent*	Pseudohomophone Translation Prime	Graphemic Control Prime
AXE	[bijl]	beil	keis	ABBEY	[abdij]	abdei	abrem
BACK	[rug]	ruch	gect	BRIDGE	[brug]	bruch	bromo
BUCKET	[emmer]	emmur	evaus	CAT	[kat]	kad	zas
CAR	[auto]	outo	bigi	CLAY	[klei]	klij	slis
CAVE	[grot]	grod	slol	COAST	[kust]	cust	cest
CHALK	[krijt]	kreit	kirie	COUGH	[kuch]	kug	gug
CHEESE	[kaas]	caas	cors	COW	[koe]	coe	cof
CHILD	[kind]	kint	siet	CRUST	[korst]	corst	carst
CITY	[stad]	stat	tont	DAY	[dag]	dach	daro
DOG	[hond]	hont	rors	DEED	[daad]	daat	dras
FORCE	[kracht]	kragt	wraki	EIGHT	[acht]	agt	ogt
FROG	[kikker]	kiccer	ijbaar	END	[eind]	eint	fing
GLUE	[lijm]	leim	lemp	FACT	[feit]	feid	flio
IRON	[ijzer]	eizer	wiber	GOAT	[geit]	gijt	grat
KING	[koning]	coning	daning	GREY	[grijs]	greis	grels
MARROW	[merg]	merch	merim	HELMET	[helm]	hellum	hellam
POPE	[paus]	pous	polm	ISLAND	[eiland]	eilant	gilang
RABBIT	[konijn]	konein	joelig	LIST	[list]	leist	luist
RIBBON	[lint]	lind	tins	MAID	[meid]	meit	slis
ROPE	[touw]	tauw	jijl	NIGHT	[nacht]	nagt	negt
SALT	[zout]	zaut	gamt	OLIVE	[olijf]	oleif	oleid
SNAIL	[slak]	slac	slau	PLEA	[pleit]	plijt	plilo
THIGH	[dij]	dei	eri	PRIZE	[prijs]	preis	prein
TOOTH	[tand]	tant	trit	SALMON	[zalm]	zallem	tallim
WASTE	[afval]	affal	agdak	SAND	[zand]	zant	jora
WIFE	[vrouw]	vrauw	alauw	SEED	[zaad]	zaat	mong
WIRE	[draad]	draat	arton	TRAIN	[trein]	trijn	trion
WOOD	[hout]	haut	spib	WAY	[weg]	wech	wino

* Translation equivalents were not displayed during the experiment.

APPENDIX B

Stimuli Experiment 2: L1 Targets with their Respective L2 Pseudohomophone Translation Primes and Graphemic Control Primes

Non-Cognates				Cognates			
Target	Translation Equivalent*	Pseudohomophone Translation Prime	Graphemic Control Prime	Target	Translation Equivalent*	Pseudohomophone Translation Prime	Graphemic Control Prime
AUTO	[car]	karr	yald	BOT	[bone]	boan	boly
BLAD	[leaf]	leav	larp	DIEF	[thief]	theef	preef
BOOM	[tree]	trea	vini	DROOM	[dream]	dreem	draim
DATUM	[date]	dait	dalt	EIK	[oak]	oack	lask
DOOLHOF	[maze]	maiz	suke	HOL	[hole]	hoal	hoil
GEVANG	[jail]	jale	zane	HUIS	[house]	howse	slask
GOLF	[wave]	waiv	shee	KAM	[comb]	kome	kimo
GRAP	[joke]	joak	wyam	KAP	[cape]	caip	jarp
GRIEP	[flue]	floo	thac	KLEI	[clay]	cley	blep
GROT	[cave]	caiv	flyn	KNIE	[knee]	knea	kned
JAS	[coat]	kote	bily	KRAAI	[crow]	croe	trym
KADER	[frame]	fraim	cralp	KRAAN	[crane]	crain	trawn
KIST	[crate]	crait	rexit	MAAT	[mate]	mait	mant
OORLOG	[war]	woar	shor	MELK	[milk]	mylk	mulk
PAD	[toad]	tode	fide	MUIS	[mouse]	mowse	moost
REM	[brake]	braik	cruso	NAAM	[name]	naim	nalm
ROOK	[smoke]	smoak	knolk	NEUS	[nose]	noze	nibe
SCHUIM	[foam]	fome	wamp	PIEK	[peek]	peec	peem
SLANG	[snake]	snaik	snabe	RIJ	[row]	wroe	prun
SLEUTEL	[key]	kea	jed	SCHAAP	[sheep]	sheap	shrap
SPEL	[game]	gaim	coxy	STEEN	[stone]	stoan	strun
TAART	[pie]	pye	gox	STOOM	[steam]	steem	starm
THUIS	[home]	hoam	hyll	TOON	[tone]	toan	toin
TOUW	[rope]	roap	joll	TREIN	[train]	trane	trune
VORM	[shape]	shaip	klegy	VLOER	[floor]	flore	plore
VROUW	[wife]	wyfe	wazz	WIJN	[wine]	wyne	wund
WINST	[gain]	gane	hend	WOORD	[word]	wurd	wrad
ZEEP	[soap]	sope	hape	ZOOL	[sole]	soal	roil

* Translation equivalents were not displayed during the experiment.

APPENDIX C

Stimuli Experiment 3: L2 Targets with their Respective L1 Pseudohomophone Associative Primes and Graphemic Control Primes

Target	Associated Translation Equivalents* [L2 – L1]	Pseudo-homo-phone Associa-tive Prime	Graphe-mic Control Prime	Target	Associated Translation Equivalents* [L2 – L1]	Pseudo-homo-phone Associa-tive Prime	Graphe-mic Control Prime
BEACH	[sand-zand]	zant	pani	MONEY	[price-prijs]	preis	krech
BIKE	[car-auto]	outo	jarp	MOUSE	[cheese-kaas]	caas	tars
BLACK	[grey-grijs]	greis	troet	NIGHT	[day-dag]	dach	jech
BOARD	[chalk-krijt]	kreit	grisp	NINE	[eight-acht]	agt	zad
BONE	[marrow-merg]	merch	verui	OIL	[olive-olijf]	oleif	oluit
BREAD	[crust-korst]	corst	horin	PATH	[road-weg]	wech	vich
BRUSH	[tooth-tand]	tant	geem	PEPPER	[salt-zout]	zaut	tuum
BUTLER	[maid-meid]	meit	gemt	QUEEN	[king-koning]	coning	ginant
CAT	[dog-hond]	hont	bost	RIVER	[bridge-brug]	bruch	truon
CHURCH	[pope-paus]	pous	zeun	SHEEP	[goat-geit]	gijt	muid
CORD	[wire-draad]	draat	drief	SLOW	[snail-slak]	slac	slir
DAY	[night-nacht]	nagt	vaur	SNEEZE	[cough-kuch]	kug	wup
DOG	[cat-kat]	kad	vid	SPINE	[back-rug]	ruch	koro
FIRE	[wood-hout]	haut	wolm	STRONG	[power-kracht]	kragt	hregt
HOLE	[cave-grot]	grod	apon	TRACK	[train-trein]	trijn	trome
HUSBAND	[wife-vrouw]	vrauw	praug	WHITE	[rabbit-konijn]	konein	ranuis
MILK	[cow-koe]	coe	eef	WIRE	[rope-touw]	tauw	pluw

* These associated translation equivalents were not displayed during the experiment. They are displayed here to illustrate the relation between the pseudohomophone associative prime and the target.

APPENDIX D

Stimuli Experiment 4: L1 Targets with their Respective L2 Pseudohomophone Associative Primes and Graphemic Control Primes

Target	Associated Translation Equivalents* [L2 – L1]	Pseudo-homo-phone Associa-tive Prime	Graphe-mic Control Prime	Target	Associated Translation Equivalents* [L2 – L1]	Pseudo-homo-phone Associa-tive Prime	Graphe-mic Control Prime
APPEL	[taart-pie]	pye	pue	MANTEL	[kap-cape]	caip	hasy
BEEN	[knie-knee]	knea	sned	OCEAAN	[golf-wave]	waiv	laky
BERG	[piek-peak]	peec	yees	PEDAAL	[rem-brake]	braik	scair
BIER	[wijn-wine]	wyne	vupe	PLAFOND	[vloer-floor]	flore	flost
BORSTEL	[kam-comb]	kome	zove	ROTS	[steen-stone]	stoan	stomi
CEL	[gevang-jail]	jale	hile	SCHOEN	[zool-sole]	soal	sorm
DRAAD	[touw-rope]	roap	reaf	SLAAP	[droom-dream]	dreem	midor
FOTO	[kader-frame]	fraim	furid	SPOOR	[trein-train]	trane	grent
GEZICHT	[neus-nose]	noze	caze	STAM	[boom-tree]	trea	tona
HEET	[stoom-steam]	steem	oteer	THUIS	[huis-house]	howse	hemsy
HOED	[jas-coat]	kote	fole	VERLIES	[winst-gain]	gane	mone
HOL	[grot-cave]	caiv	sepa	VREDE	[oorlog-war]	woar	plur
HUIS	[thuis-home]	hoam	hacy	VRIEND	[maat-mate]	mait	coit
KAT	[muis-mouse]	mowse	scoze	VUUR	[rook-smoke]	smoak	ebaga
KIKKER	[pad-toad]	tode	cune	WOL	[schaap-sheep]	sheap	vunge
KOE	[melk-milk]	mylk	zurk	ZEEP	[schuim-foam]	fome	dige
KOLOM	[rij-row]	wroe	apos	ZIEK	[griep-flu]	floo	turg
LACH	[grap-joke]	joak	piam	ZIN	[woord-word]	wurd	pulm

* These associated translation equivalents were not displayed during the experiment. They are displayed here to illustrate the relation between the pseudohomophone associative prime and the target.

APPENDIX E

Stimuli Experiment 5 and 6: L2/L1 Targets with their intralingual homophones of their respective L1/L2 Translation Equivalents and Graphemic Control Primes

L2 Targets/Primes (Experiment 5)				**L1 Targets/Primes (Experiment 6)**			
L2 Target	L1 Translation Equivalent*	L2 Homophone Prime	L2 Graphemic Control Prime	L1 Target	L2 Translation Equivalent*	L1 Homophone Prime	L1 Graphemic Control Prime
ARROW	[pijl]	pale	camp	AAS	[ace]	ijs	les
AXE	[bijl]	bail	cart	BAAI	[bay]	bij	bil
BEE	[bij]	bay	bow	BAAS	[boss]	bos	bes
CHALK	[krijt]	crate	candy	BORG	[bail]	bijl	beek
CORNER	[hoek]	hook	foot	DAG	[day]	dij	dek
CORPSE	[lijk]	lake	bite	DIJ	[thigh]	taai	klei
COURAGE	[moed]	mood	book	HAAK	[hook]	hoek	heks
FACT	[feit]	fate	fast	HOOI	[hay]	hei	hik
GLUE	[lijm]	lame	lobe	HUMEUR	[mood]	moed	melk
GOLD	[goud]	goat	gone	KAMER	[room]	roem	riem
JOURNEY	[reis]	raise	rifle	KLOP	[knock]	nok	wok
LEAK	[lek]	lack	lark	LOF	[praise]	prijs	maand
LINE	[lijn]	lane	lone	LUS	[loop]	loep	lijm
MILE	[mijl]	male	mule	MEER	[lake]	lijk	zout
PLEA	[pleit]	plate	plane	PLAAT	[plate]	pleit	plint
PRICE	[prijs]	praise	priest	POORT	[gate]	geit	kant
ROW	[rij]	ray	rub	PROOI	[prey]	prei	prik
SCREW	[vijs]	vase	fuse	RIJVAK	[lane]	lijn	tijd
SONG	[lied]	lead	pick	RUIMTE	[space]	spijs	klink
SQUARE	[plein]	plane	whale	STAAL	[steel]	stiel	stijl
THIGH	[dij]	day	new	STAART	[tail]	teil	toog
TIME	[tijd]	tide	tile	STRAAL	[ray]	rij	ruk
TRAIT	[trek]	track	trace	WEG	[way]	wei	web

* These intermediate translation equivalents were not displayed during the experiment. They are displayed here to illustrate the relation between the intralingual (homophone of the translation equivalent) prime and the target.

CHAPTER 5
THE SIZE OF THE CROSS-LINGUAL MASKED PHONOLOGICAL PRIMING EFFECT DOES NOT DEPEND ON SECOND LANGUAGE PROFICIENCY

***Experimental Psychology, 2004, 51, 1-9*[1, 2]**

Using a masked phonological priming paradigm, Brysbaert, Van Dyck and Van de Poel (1999) showed that Dutch-French bilinguals perform better at identifying tachistoscopically presented L2 words (e.g. *oui* [yes]) when those words are primed by L1 words or nonwords that are homophonic to the L2 target word according to the L1 grapheme-phoneme conversion rules (e.g. *wie* [who]). They noted that this priming effect was smaller for balanced bilinguals than for less proficient bilinguals, although the interaction failed to reach significance. Findings of Gollan, Forster and Frost (1997) suggest that this could be attributed to a greater reliance on phonology in L2 reading, caused by a smaller proficiency in this language. However, in this study we show that the Dutch-French cross-lingual phonological priming effect is equally large for perfectly balanced and less proficient bilinguals. Our findings are in line with more recent work of Van Wijnendaele and Brysbaert (2002).

---

[1] This paper was co-authored by Kevin Diependaele, Denis Drieghe and Marc Brysbaert.

[2] The authors would like to thank Wouter De Baene, Christoph Klauer, Wido La Heij, Arnaud Szmalec and an anonymous reviewer for comments on an earlier draft of this paper.

INTRODUCTION

For many years, it has been assumed that the lexicons of every language mastered by a bilingual person are separate, autonomous systems. In older models of bilingual brain organization, such as the three models of Weinreich (1953), both languages are completely divided at the lexical level, while shared representations between languages may exist at the semantic level. This assumption was also made in more recent models of bilingualism, such as the word association model, the concept mediation model (Potter, So, Von Eckardt, & Feldman, 1984) and the revised hierarchical model of Kroll and Stewart (1994). This hypothesis is supported for example by the existence of double dissociations between both languages in bilingual aphasic patients (Fabbro, 1999).

However, recently there is a growing body of evidence suggesting that lexical representations of both languages may be situated within a unitary system, or that lexical selection is at least a relatively late process in visual word recognition. A somewhat older study which already pointed in that direction is that of Nas (1983). He showed that Dutch-English bilinguals performing an English lexical decision task rejected Dutch words significantly slower than control words. The same was true for English nonwords (e.g. *snay*) which are homophones of existing Dutch words (e.g. *snee*, translated *cut*) according to English grapheme-to-phoneme conversion rules. However, using Spanish-English bilinguals, Scarborough, Gerard and Cortese (1984) did not replicate the findings of Nas (1983). Grainger (1993) argued that the effect was absent in the latter study because the orthographic similarity between Dutch and English (two Germanic languages) is much larger than between Spanish and English. Consequently, participants were more likely to have performed the lexical decision task using nonlexical (e.g. orthographic) characteristics of the target words.

More recently, Bijeljac-Babic, Biardeau and Grainger (1997) found that recognition of low-frequency target words by French-English bilinguals is inhibited not only by intra-lingual, but also by cross-lingual high-frequency orthographic neighbour primes (e.g. recognition of the French word *amont* is more difficult after masked presentation of the English prime *among* than after the control word *drive*). Another group of studies favouring the integrated lexicon hypothesis makes use of interlingual homographs (words which exist in both languages but have different meanings, e.g. the English word *room* means *cream* in Dutch). De Groot, Delmaar and Lupker (2000) for example, showed that the processing of interlingual homographs in a translation recognition task was inhibited compared to the processing of matched control words. This was especially the case when the homograph reading to be selected was the less frequent of the two homograph's readings. Dijkstra, Timmermans and Schriefers (2000) showed that such frequency dependent inhibitory effects of interlingual homographs are also present in tasks which do not explicitly require simultaneous activation of both language systems (this is the case in a translation recognition paradigm as De Groot et al. used). This shows that the presence of both languages in the experimental stimuli is not a necessary condition to find cross-language lexical interactions. Moreover, Van Hell and Dijkstra (2002) recently showed that L2 and even L3 lexical knowledge influences L1 lexical access in an exclusive native language context, using a L1 lexical decision task with Dutch – English – French trilinguals. Even though no L2 of L3 words (e.g. homographs, Dijkstra et al., 2000, see earlier) were present in the experiment, they found that L1 lexical decision is faster for L2 and L3 near-cognates (i.e. translation equivalents which are nearly orthographically identical, e.g. *brood* – *bread*) than for control words. Hence, this strongly suggests that L1 lexical activation is influenced by activation in the lexical representations of L2 and L3 words. For a more comprehensive overview of studies favouring the unitary lexical system view, we refer to Dijkstra and Van Heuven (1998; see also Brysbaert, 1998). For the present study, it is only important to conclude that several recent studies have provided evidence against an early lexical selection mechanism.

Based on this body of evidence and on the claim that visual word recognition implies automatic, prelexical phonological coding (e.g. Van Orden, 1987; see Frost, 1998, for a recent review), Brysbaert et al. (1999) reasoned that it is very likely that such an automatic (not strategically controlled) grapheme-to-phoneme conversion occurs for all grapheme-phoneme correspondences mastered by bilinguals. This conversion takes place before a language selection mechanism gets involved in the word recognition process. This is compatible with an earlier study of Doctor and Klein (1992) with English-Afrikaans bilinguals. They found that interlingual homophones (words which share the same pronunciation, but have a different spelling, e.g. *lake* and *lyk* [corps]) are processed slower and less accurately than control words in a lexical decision task. To investigate this hypothesis more directly, Brysbaert et al. made use of the masked phonological priming effect, which was first reported by Humphreys, Evett and Taylor (1982). In this study, they showed that recognition of a tachistoscopically presented target word (e.g. *mail*) is facilitated by presentation of a masked homophonic prime (e.g. *male*) relatively to a graphemic control prime (e.g. *mall*). The difference between recognition ratios in those two conditions will be referred to as the (net) phonological priming effect from this point on. Note that this priming effect can not easily be attributed to strategic factors, since participants are unable to perform above chance in deciding whether the prime was a word or not, even when they are asked to try to identify the prime (e.g. (e.g. Forster & Davis, 1984).

In a first experiment, Brysbaert et al. (1999) used a bilingual version of this paradigm, using French target words and Dutch primes: the target words (e.g. *nez,* translation *nose*) were presented tachistoscopically preceded by either homophonic primes (e.g. *nee*, translated *no*, sounds like the French word *nez*), graphemic control primes (e.g. *nek*, translated *neck*), or unrelated primes (e.g. *oud,* translated *old*). Note that the L1 homophonic primes were only homophonic with the L2 target word according to L1 (Dutch) grapheme-to-phoneme conversion rules. They found that target recognition was equally well in the homophonic and graphemic control condition for French monolinguals, but not for Dutch-French bilinguals. The latter performed significantly better after seeing the homophonic prime, than after seeing the graphemic control prime. To counter the criticism that this effect could be due to interactions within the bilingual's input lexicon, or between two language-dependent input lexicons, these findings were replicated with Dutch nonwords in a second experiment (e.g. a French target *pour* [translation *for*], with *poer*, *poir* and *dalk* as respectively Dutch homophonic, graphemic control and unrelated nonword primes). These results are evidence for automatic, prelexical and language-independent phonological coding of orthographic stimuli. Similarly, recent research (Van Wijnendaele & Brysbaert, 2002) offers further support for strong phonological models of word recognition (e.g. Van Orden & Goldinger, 1994; Frost, 1998). Using Dutch-French bilinguals, they found that it is also possible to prime L1 words (e.g. *wie* [who]) with L2 homophonic primes (e.g. *oui* [yes[). Hence, it is not only the case that word forms are automatically phonologically coded according to L1 grapheme-to-phoneme conversion rules. The same applies for L2 grapheme-to-phoneme rules, even when performing a task in L1 (see further in this introduction). Such a result can not be easily explained by traditional dual-route models of visual word recognition (e.g. Coltheart, 1978; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001). A more detailed discussion of this study and further interpretation of the results within these models is beyond the scope of this paper (see Van Wijnendaele & Brysbaert, 2002).

A less discussed and analysed though very intriguing aspect of the Brysbaert et al. (1999) study is the observation that the cross-lingual phonological priming effect was smaller for participants who learned French from birth than for those who started to learn French around the age of 10. For the first group, the difference between the proportions of correctly identified targets in the homophonic and graphemic control condition was .00 (.03 en -.03 for respectively Experiments 1 and 2). For the late learners, the effect was .10 (respectively .09 and .11). However, the interaction of the cross-lingual phonological priming effect with second language proficiency failed to reach significance (no *p* values mentioned). It should be noted though that this issue was not of primary concern for Brysbaert et al., and that only eight out of 40 (Experiment 1) and five out of 30 (Experiment 2) participants were balanced bilinguals. Hence, their study was not optimally designed to find such an interaction. In this study, we will focus on this topic, and we will therefore present some data of larger groups of perfectly balanced and other bilinguals performing the task used in Brysbaert et al. This allows us to determine whether the finding of Brysbaert et al. may be due to the use of a limited sample of balanced bilinguals.

Finding this interaction effect would be in line with results found in a Hebrew-English masked translation study by Gollan, Forster and Frost (1997). They found that it is possible to prime L2 targets with L1 translation primes while the priming effect from L2 primes to L1 targets was much weaker and not consistent. Because their primes contained both non-cognates (semantic overlap) and cognates (semantically and phonologically, but not orthographically overlapping as Hebrew and English have different scripts), they attributed this observation to the fact that L2 reading may rely more on phonology than L1 reading. It is indeed plausible to assume that L2 target recognition is more susceptible to phonologically similar primes than L1 target recognition if this explanation is correct. Note that such cross-language priming asymmetries have also been reported more recently by Jiang and Forster (2001), though they explained this finding differently.

Gollan et al. (1997) also stated that this overreliance on phonology in L2 reading is caused by a smaller L2 proficiency relatively to L1. This hypothesis is congruent with their observation of a larger cognate effect for less proficient than for more balanced bilinguals: L2 target recognition was facilitated by presentation of L1 cognate primes relatively to L1 non-cognate (phonologically dissimilar) primes , and this facilitation effect was greater for less proficient bilinguals. Hence, the phonological overlap between the L1 cognate prime and the L2 target was of greater importance for less proficient bilinguals, suggesting a larger reliance on phonological codes. Thus, on the basis of these findings, one would also predict a negative correlation between the cross-lingual phonological priming effect obtained by Brysbaert et al. (1999) and L2 proficiency in our study: perfectly balanced bilinguals will rely less on phonology than other bilinguals when processing L2 target words. Therefore, L2 target recognition will be less influenced by presentation of homophonic L1 primes and the cross-lingual phonological priming effect will be smaller for balanced bilinguals, as found by Brysbaert et al..

However, while an interaction between L2 proficiency and the cross-language phonological priming effect may be expected on the basis of the Gollan et al. (1997) study, recent findings suggest the contrary: as noted earlier, Van Wijnendaele and Brysbaert (2002) found in a Dutch-French study that it is also possible to prime L1 targets with homophonic L2 primes. This priming effect was of the same magnitude as the cross-lingual phonological priming effect from L1 to L2 (Brysbaert et al., 1999). Moreover, both priming effects were not related to differences in word naming latencies between L1 and L2 (r = -.17, p > .10), a variable believed to reflect language proficiency (Van Wijnendaele & Brysbaert, 2002; La Heij, Hooglander, Kerling, & Van der Velden, 1996; Kroll & Stewart, 1994). In addition, no evidence has been found in this study for an overreliance on phonology in L2 reading, as hypothesized by Gollan et al.. On the contrary, there was a larger word-frequency effect for L2 word naming than for L1, suggesting less non-lexical grapheme-to-phoneme conversions in L2 reading.

EXPERIMENT

METHOD

Participants. The participants consisted of two groups of Dutch-French bilinguals. The first group were 25 students at Ghent University, who participated for course requirements. They had started to learn French in a scholastic setting around the age of 9-10. The second group were 25 balanced bilinguals who learned French from birth and who grew up in a bilingual environment (e.g. having a Dutch speaking mother and a French speaking father). Ten of them were from the same population as mentioned above. The other 15 participants participated voluntarily after responding to an e-mail announcement. All participants from the second group reported regular use of both French and Dutch in their domestic environment at the time of the experiment. All participants completed a questionnaire assessing their L1 and L2 proficiency (see further).

Stimulus Materials. The stimuli (see the Appendix) consisted of the 30 French target words matched with three types of Dutch primes collected by Brysbaert et al. (1999). Homophonic Dutch primes had the same pronunciation (according to Dutch grapheme-to-phoneme conversion rules) as the corresponding French target word (e.g. *kraan* – CRANE; translation *tap* – SKULL). Graphemic control primes had a different pronunciation, but had those letters in common with the homophonic prime that the latter shared with the target in the same letter position (e.g. *graan* – CRANE; translation *grain* – SKULL). Finally, unrelated control primes had neither letters nor sounds in common with the target (e.g. *stoom* – SKULL; translation *steam* – SKULL). This type of control prime (Berent & Perfetti, 1995) is included to check the effectiveness of the priming procedure in case differences between the first two prime conditions would be absent. There was no semantic overlap between the primes and the target, and care was also taken that no Dutch prime was also an existing French word, or was homophonic to the target word according to French grapheme-to-phoneme conversion rules. Also, the log frequency of the three Dutch primes was matched (based on the CELEX counts, Baayen, Piepenbrock, & Van Rijn, 1993). The mean printed frequency of the target words was 366 per million (*Trésor de la Langue Française,* 1971).

Procedure. The same procedure was used as in Brysbaert et al. (1999). Participants were tested in small groups. Care was taken that they were placed sufficiently far from each other. It was not possible to see the computer screen of another participant. First, the instructions were presented on the screen in French. They mentioned that five practice trials and 30 experimental trials would follow. At the beginning of each trial, two vertical lines appeared as a fixation point in the center of the screen. Participants were also instructed to press the space bar to continue with the next trial. Five hundred milliseconds after this keypress, a forward mask consisting of seven hash-marks (#######) was presented with the second sign at the place of the gap between the two vertical lines. This mask stayed on the screen for another 500 ms, and was followed by a prime for 42 ms, a target word for 42 ms and a postmask consisting of seven horizontally aligned capital Xs (*XXXXXXX*). This mask remained visible until the end of the trial. The timing of the stimulus presentation was controlled using software routines published by Bovens and Brysbaert (1990). The prime appeared in lowercase letters, unlike the target which appeared in uppercase letters (for this reason, Xs were used as a more effective postmask). Both primes and targets were presented at the optimal viewing position (i.e., the second letter always appeared between the two vertical lines, e.g. Brysbaert, Vitu, & Schroyens, 1996). Participants were warned that on each trial a French word in uppercase letters would appear on the screen, and they were instructed to identify the word and type it in. There was no mentioning of the Dutch prime words. The letters typed in by the participants were automatically converted on the screen into uppercase letters to avoid the need to type accent marks. Each participant received a random permutation of the 30

Dutch-French stimuli. Therefore, each target word was only presented once, with one type of prime stimulus (Latin-square design).

Finally, all participants also completed a questionnaire, assessing their self-reported L1 and L2 reading, speaking, writing and general proficiency level on a seven-point Likert scale ranging from ‘very bad’ to ‘very good’. In addition, the questionnaire contained some general questions regarding the participants’ history of L2 acquisition (e.g. setting, age, etc.).

RESULTS

Balanced and unbalanced bilinguals differed significantly with respect to their reported L2 speaking proficiency (respective means were $M = 5.84$ and $M = 3.88$, $F(1, 48) = 52.39$, $MSE = .917$, $p < .001$), writing proficiency ($M = 5.56$ and $M = 3.60$, $F(1, 48) = 45.95$, $MSE = 1.045$, $p < .001$) and reading proficiency ($M = 6.12$ and $M = 4.40$, $F(1, 48) = 36.98$, $MSE = .638$, $p < .001$). Balanced bilinguals also reported significantly higher general L2 proficiency, $M = 5.95$ and $M = 3.88$, $F(1, 48) = 52.02$, $MSE = .635$, $p < .001$. Both groups did not differ with respect to L1 speaking, writing, reading and general proficiency. Accordingly, the age at which participants reported to have encountered their first L2 word was significantly lower for balanced bilinguals ($M = 1.92$) than for unbalanced bilinguals ($M = 8.96$), $F(1, 48) = 397.55$, $MSE = 1.558$, $p < .001$. Consequently, the balanced bilinguals also had significantly more years of L2 experience ($M = 21.04$ vs. $M = 11.56$ years), $F(1, 48) = 48.38$, $MSE = 9.023$, $p < .001$.

Probabilities of correct target word identification as a function of prime type and bilingual group are displayed in Table 1. ANOVAs were run with L2 proficiency (balanced versus other bilinguals), prime type (homophonic, graphemic and control) and Latin-square group as independent variables. The latter variable was included to correct for the possibly deflated power of the design due to random fluctuations between the participants or between the stimuli allocated to the different cells. This has shown to be a good solution when analyzing Latin-square designs with relatively few observations in the different cells (Pollatsek & Well, 1995).

The main effect of Prime Type was significant both in the analysis by participants and by items, $F_1(2, 88) = 14.34$, $MSE = .0094$, $p < .01$, $F_2(2, 54) = 4.89$, $MSE = .0351$, $p < .01$. Because we had precise predictions concerning the phonological priming effect at the onset of the study, we could legitimately run a planned comparisons analysis. This showed a significant difference between the homophonic and the graphemic control condition, both in the analysis by participants and items, $F_1(1, 44) = 12.57$, $MSE = .0096$, $p < .001$, $F_1(1, 27) = 5.34$, $MSE = .0287$, $p < .03$. There were no significant main effects of Latin-square group (both Fs < 1) and L2 proficiency ($F_1 < 1$, $F_2(1, 27) = 1.55$, $p > .20$).

Prime Type	Example	Less Proficient Bilinguals	Highly Proficient Bilinguals	Mixed Bilinguals (Brysbaert et al., 1999)
Homophonic	kraan – CRANE	23.3%	26.4%	30%
Graphemic Control	graan – CRANE	16.4%	19.0%	23%
Unrelated Control	stoom – CRANE	14.9%	13.9%	17%
Net Phonological Priming Effect		6.9%	7.4%	7%

Table 1. Probabilities (%) of Correct Target Word Identification as a Function of L2 Proficiency and Prime Type

Most importantly, no significant interaction was found between L2 proficiency and primetype, Fs < 1 ($MSE_1 = .0094$, $MSE_2 = .0091$). Also, a planned comparison of the interaction between L2 proficiency and the two primetype conditions involved in the phonological priming effect was not significant, both Fs < 1 ($MSE_1 = .0096$, $MSE_2 = .0122$). To evaluate the strength of this finding, we analyzed the power of this test in our design using the procedure of the MorePower program developed by Campbell and Thompson (2002). Because of the quite large number of participants and the rather small variance in the phonological priming effect, the design had a .805 power to detect the average net phonological effect difference between balanced and unbalanced bilinguals reported by Brysbaert, Van Dyck and Van de Poel (1999) (one-tailed), which is higher than the generally accepted .80 power level. There was even a very small trend towards a larger phonological priming effect for balanced bilinguals (7.4%) compared to other bilinguals (6.9%), rather than a smaller (or absent) effect.

Finally, whereas Table 1 suggests a larger difference between the graphemic and the unrelated control condition for balanced (5.1%) than for other (1.5%) bilinguals, a planned comparison showed that this interaction was by no means significant, $F_1 < 1$, $MSE = .0088$, $F_2(1, 27) = 1.13$, $MSE = .0086$, $p > .29$.

DISCUSSION

The results of the experiment are quite clear: although mean target recognition rate was somewhat lower than in the study of Brysbaert et al. (1999), we succeeded in replicating the Dutch-French cross-lingual phonological priming effect[3]. Moreover, the effect we found was almost of the exact same size (it was 7.1% in our study, while it was 7.0% in Brysbaert et al.). In terms of statistical reliability (especially in the analysis by materials), the effect was somewhat stronger in this study, probably because of the larger number of participants.

Most importantly, this cross-lingual phonological priming effect did not interact with L2 proficiency, contrary to our predictions based on the findings of Brysbaert et al. (1999) and Gollan et al. (1997). Hence, it seems

---

[3] Note again that the absence of such an effect in a monolingual French control group (using the same stimuli, Brysbaert et al., 1999, Experiment 1) rules out the possibility that the origin of the phonological priming effect lies within the (unevitable) orthographic overlap of the homophonic primes with the target.

that the (not significantly) smaller priming effect for balanced bilinguals found by Brysbaert et al. (a .10 difference averaged over experiments) was indeed due to random fluctuations within the limited sample (n = 8 and n = 5, Experiments 1 and 2) of balanced bilinguals they used. Our findings are also not completely compatible with Gollan et al.. They reported an interaction between L2 proficiency and the cross-lingual cognate effect. The phonological overlap between a L1 cognate prime and a L2 target (Hebrew-English cognates are semantically and phonologically, but not orthographically similar) was of less importance for highly proficient bilinguals than for less proficient bilinguals. This suggests a negative correlation between L2 proficiency and the importance of phonological codes in L2 word reading. This hypothesis has not been confirmed in our experiment: the cross-lingual phonological priming effect was equally large for both groups (7.4% versus 6.9%: there was even a small trend in the opposite direction of predictions based on the results of Gollan et al. with a larger effect for balanced bilinguals). This would not have been the case if phonological codes are less important for L2 word recognition in perfectly balanced bilinguals.

However, our results are in line with more recent cross-lingual priming research of Van Wijnendaele and Brysbaert (2002): they found (unlike Gollan et al., 1997; Jiang & Forster, 2001) in a Dutch-French study that it is also possible to prime L1 targets with homophonic L2 primes. This priming effect was not smaller than the cross-lingual phonological priming effect from L1 to L2 (Brysbaert et al., 1999), which is against Gollan et al.'s hypothesis of an overreliance on phonology in L2 reading, for this hypothesis implies L2 target recognition to be more influenced by homophonic primes than L1 target recognition. Most importantly, both cross-lingual priming effects were also not related to differences in word naming latencies between L1 and L2, which were used to assess L2 proficiency.

In this view, it may be plausible to attribute the priming asymmetry (i.e. forward priming from L1 to L2, but not backward from L2 to L1) observed by Gollan et al. (see also Jiang, 1999, who replicated this asymmetry in Chinese-English bilinguals), not to a greater reliance on phonology in L2 reading relative to L1 reading, but to the fact that Hebrew (their L1) and English (L2) have different alphabets and therefore share little, if any, orthographic features (Grainger, 1993; Brysbaert, 2003). This is clearly not the case for Dutch and French which are orthographically more similar, and which are also much more consistent relative to each other as grapheme to conversion rules are concerned. This has probably facilitated transfer of phonological activation between languages, although our present findings do not allow to make strong claims about this issue. However, it may be interesting to note that we recently found both forward and backward translation priming in Dutch-English bilinguals using a lexical decision task (Schoonbaert, Duyck, & Brysbaert, 2003), whereas only L1 to L2 priming was reported by Jiang and Forster (2001), again using two languages which have different alphabets (i.e. Chinese and English). Finally, one might also argue that phonological codes are more important for the L2 perceptual identification task used in this study than for the L2 lexical decision task used by Gollan et al. (1997). However, Grainger and Ferrand (1996) compared these two tasks directly with the same set of stimuli and found a robust (intra-lingual) phonological masked priming effect with both tasks. Also, Kim and Davis (2003) recently found a phonological cross-language priming effect with Korean-English bilinguals using a lexical decision task (although this 18ms effect was only significant in a one-tailed test).

The present findings offer further evidence against the existence of two independent lexical language systems, since those models are unable to explain cross-language interactions at such an early stage of visual word recognition (see also Bijeljac-Babic et al., 1997). In order to avoid between-language confusion, inhibition of an irrelevant language system is likely to occur at some point, but this and other mentioned evidence suggest that this stage occurs relatively late in visual word recognition. An example of a powerful model which does not postulate language-specific access to the mental lexicon is the Bilingual Interactive Activation (BIA) model of Dijkstra and Van Heuven (1998). This is an extension of the well-known Interactive Activation model for monolingual word recognition (e.g. McClelland & Rumelhart, 1981), in which a top-down activation flow of language nodes to word nodes is made possible to account for language inhibition and facilitation effects on the word-node level. Hence, the monolingual model has been extended by a) adding language nodes (supplementary to word, letter and feature nodes) and b) inclusion of all L2 words into a unitary word-level system. This model implies that word recognition processes are initially non-selective (though top-down language influences may exist), since word activation is affected by competing items from both languages (e.g. Van Heuven, Dijkstra, & Grainger, 1998). Note that in more recent versions of the BIA model (see Dijkstra and Van Heuven, 2002, for a description of the BIA+ model), all top-down connections have been removed. Instead, effects of language context and stimulus list composition are dealt with at the task schema level, which only receives input from the (fundamentally language non-selective) word identification system. In this architecture, decision criteria, in a lexical decision task for example, can change as a function of stimulus list composition (e.g. Dijkstra et al., 2000), without assuming that such top-down factors influence activation in the lexical representations itself. Unlike the older BIA model, the BIA+ model also contains semantic and phonological representations, although these have not been implemented yet. It will be very interesting to see whether this model will be able to cope with the cross-lingual phonological priming effect. The present study and the findings of Brysbaert et al. (1999) and Van Wijnendaele and Brysbaert (2002) strongly suggest that activation of these phonological representations will also have to be fundamentally language non-selective, just as for lexical representations. In this view, we would also like to note that the phonological priming effect was equally strong from L2 to L1 than in the other direction (Van Wijnendaele & Brysbaert, 2002). This is not entirely compatible with the temporal delay assumption of the BIA+ model, which states that L2 phonological and semantic representations are delayed in activation relative to L1 codes (the same might be true for semantic representations, e.g. see Duyck & Brysbaert, 2002). As a more detailed discussion of this issue is beyond the scope of this paper, we refer the interested reader to Brysbaert, Van Wijnendaele and Duyck (2002).

Other models of bilingual word recognition in which some degree of interconnectedness of both languages is assumed (although to a lesser extent), such as the Bilingual Model of Lexical Access (BIMOLA) of Grosjean (1988; 1997), are less compatible with the cross-lingual phonological priming effect than the BIA model. In BIMOLA, there are two independent language networks (features, phonemes, words, etc.) which are both activated to some degree, depending on higher linguistic (e.g. textual context) information. Both systems are interconnected by means of a subset of neural connections from which bilinguals are able to draw elements of both languages, supplementary to the subset of neural connections for each separate language. Hence, this model can only predict interactions between two languages at such an early stage when higher linguistic information triggers activation in and between both language networks. This is not self-evident in a French target recognition task when participants are not aware of the presence of Dutch primes (e.g. Forster & Davis, 1984). In that case, the model (operating in a monolingual language mode) would not predict much influence from the weakly activated Dutch language system on the more strongly activated French language network. It should be noted though that there has recently been some evidence (Jared & Kroll, 2001) for Grosjean's (1988; 1997; 2001) claim that the task environment becomes functionally bilingual if the participant expects the experiment in which he or she is about to participate is likely to be using both languages, even if only materials in a single language are presented. This may have been the case since our participants were recruited based on their bilingual history.

In conclusion, it can be stated that our results offer further support for a strong phonological view on word recognition (e.g. Van Orden, 1987; Frost, 1998; Dijkstra, Grainger, & Van Heuven, 1999; for a recent and more detailed discussion, see Van Wijnendaele & Brysbaert, 2002): visual input triggers automatic phonological activation, and this occurs for all grapheme-to-phoneme correspondences mastered by a bilingual (see also Doctor & Klein, 1992). Moreover, contrary to Gollan et al. (1997), this process does not interact with L2 proficiency: phonology plays a crucial role in L2 word recognition, even in perfectly balanced bilinguals, as shown by the relatively large cross-lingual phonological priming effect in this group.

REFERENCES

Baayen, R., Piepenbrock, R., & Van Rijn, H. (1993). *The CELEX lexical database*. [CD-ROM]. Philadelphia: University of Pennsylvania: Linguistic Data Consortium.

Berent, I., & Perfetti, C. A. (1995). A Rose Is a Reez - the 2-Cycles Model of Phonology Assembly in Reading English. *Psychological Review, 102*(1) , 146-184.

Bijeljac-Babic, R., Biardeau, A., & Grainger, J. (1997). Masked Orthographic Priming in Bilingual Word Recognition. *Memory & Cognition, 25*(4), 447-457.

Bovens, N., & Brysbaert, M. (1990). Ibm-Pc/Xt/at and Ps/2 Turbo Pascal Timing With Extended Resolution. *Behavior Research Methods Instruments & Computers, 22*(3), 332-334.

Brysbaert, M. (1998). Word Recognition in Bilinguals: Evidence Against the Existence of Two Separate Lexicons. *Psychologica Belgica, 38*(3-4), 163-175.

Brysbaert, M. (2003). Bilingual Visual Word Recognition: Evidence from Masked Phonological Priming. In Kinoshita, S. and Lupker, S. J. (Eds.), *Masked priming: State of the Art.* (pp. pp. 323-343). Hove, UK: Psychology Press.

Brysbaert, M., Van Dyck, G., & Van De Poel, M. (1999). Visual Word Recognition in Bilinguals: Evidence From Masked Phonological Priming. *Journal of Experimental Psychology: Human Perception and Performance, 25*(1), 137-148.

Brysbaert, M., Van Wijnendaele, I., & Duyck, W. (2002). On the temporal delay assumption and the impact of non-linguistic context effects. *Bilingualism: Language and Cognition, 5*(3), 199-201.

Brysbaert, M., Vitu, F., & Schroyens, W. (1996). The Right Visual Field Advantage and the Optimal Viewing Position Effect: on the Relation Between Foveal and Parafoveal Word Recognition. *Neuropsychology, 10*(3), 385-395.

Campbell, J. I. D., & Thompson, V. A. (2002). More Power to You: Simple Power Calculations for Treatment Effects With One Degree of Freedom. *Behavior Research Methods Instruments & Computers, 34*(3), 332-337.

Coltheart, M. (1978). Lexical Access In Simple Reading Tasks. In Underwood, G. (Ed.), *Strategies Of Information Processing.* (pp. 151-216). London: Academic Press.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). Drc: a Dual Route Cascaded Model of Visual Word Recognition and Reading Aloud. *Psychological Review, 108*(1), 204-256.

De Groot, A. M. B., Delmaar, P., & Lupker, S. J. (2000). The Processing of Interlexical Homographs in Translation Recognition and Lexical Decision: Support for Non-Selective Access to Bilingual Memory. *Quarterly Journal of Experimental Psychology, 53A*(2), 397-428.

Dijkstra, A., & Van Heuven, W. J. B. (1998). The BIA Model and Bilingual Word Recognition. In Grainger, J. and Jacobs, A. (Eds.), *Localist Connectionist Approaches to Human Cognition.* (pp. 189-225). Hillsdale, NJ: Erlbaum.

Dijkstra, T., Grainger, J., & Van Heuven, W. J. B. (1999). Recognition of Cognates and Interlingual Homographs: the Neglected Role of Phonology. *Journal of Memory and Language, 41*(4), 496-518.

Dijkstra, T., Timmermans, M., & Schriefers, H. (2000). On Being Blinded by Your Other Language: Effects of Task Demands on Interlingual Homograph Recognition. *Journal of Memory and Language, 42*(4), 445-464.

Doctor, E. A., & Klein, D. (1992). A Model of Bilingual Word Recognition and Dyslexia - Its Use in Assessment Diagnosis and Remediation. *International Journal of Psychology, 27*(3-4), 392

Duyck, W., & Brysbaert, M. (2002). What Number Translation Studies Can Teach Us About the Lexico-Semantic Organisation in Bilinguals. *Psychologica Belgica, 42*(3), 151-175.

Fabbro, F. (1999). *The Neurolinguistics Of Bilingualism: An Introduction*. Psychology Press.

Forster, K. I., & Davis, C. (1984). Repetition Priming and Frequency Attenuation in Lexical Access. *Journal of Experimental Psychology: Learning, Memory and Cognition, 10*(4), 680-698.

Frost, R. (1998). Toward a Strong Phonological Theory of Visual Word Recognition: True Issues and False Trails. *Psychological Bulletin, 123*(1), 71-99.

Gollan, T. H., Forster, K. I., & Frost, R. (1997). Translation Priming With Different Scripts: Masked Priming With Cognates and Noncognates in Hebrew-English Bilinguals. *Journal of Experimental Psychology: Learning, Memory and Cognition, 23*(5), 1122-1139.

Grainger, J. (1993). Visual Word Recognition in Bilinguals. In Schreuder, R. and Weltens, B. (Eds.), *The Bilingual Lexicon.* (pp. 11-26). Amsterdam: John Benjamins.

Grainger, J., & Ferrand, L. (1996). Masked Orthographic and Phonological Priming in Visual Word Recognition and Naming: Cross-Task Comparisons. *Journal of Memory and Language, 35*(5), 623-647.

Grosjean, F. (1988). Exploring the Recognition of Guest Words in Bilingual Speech. *Language and Cognitive Processes, 3*, 233-274.

Grosjean, F. (1997). Processing Mixed Language: Issues, Findings, and Models. In De Groot, A. M. B. and Kroll, J. F. (Eds.), *Tutorials In Bilingualism: Psycholinguistic Perspectives.* (pp. 225-254). Mahwah, NJ: Lawrence Erlbaum Publishers.

Grosjean, F. (2001). The Bilingual's Language Modes. In Nicol, J. L. (Ed.), *One Mind, Two Languages: Bilingual Language Processing.* (pp. Oxford: Blackwell.

Humphreys, G. W., Evett, L. J., & Taylor, D. E. (1982). Automatic Phonological Priming in Visual Word Recognition. *Memory & Cognition, 10*(6), 576-590.

Jared, D., & Kroll, J. F. (2001). Do Bilinguals Activate Phonological Representations in One or Both of Their Languages When Naming Words? *Journal of Memory and Language, 44*(1), 2-31.

Jiang, N. (1999). Testing Processing Explanations for the Asymmetry in Masked Cross-Language Priming. *Bilingualism: Language and Cognition, 2*, 59-75.

Jiang, N., & Forster, K. I. (2001). Cross-Language Priming Asymmetries in Lexical Decision and Episodic Recognition. *Journal of Memory and Language, 44*(1), 32-51.

Kim, J., & Davis, C. (2003). Task Effects in Masked Cross-script Translation and Phonological Priming. *Journal of Memory and Language, 49*(4), 484-499.

Kroll, J. F., & Stewart, E. (1994). Category Interference in Translation and Picture Naming - Evidence for Asymmetric Connections Between Bilingual Memory Representations. *Journal of Memory and Language, 33*(2), 149-174.

La Heij, W., Hooglander, A., Kerling, R., & Van der Velden, E. (1996). Nonverbal Context Effects in Forward and Backward Word Translation: Evidence for Concept Mediation. *Journal of Memory and Language, 35*(5), 648-665.

McClelland, J. L., & Rumelhart, D. E. (1981). An Interactive Activation Model of Context Effects in Letter Perception .1. An Account of Basic Findings. *Psychological Review, 88*(5), 375-407.

Nas, G. (1983). Visual Word Recognition in Bilinguals - Evidence for a Cooperation Between Visual and Sound Based Codes During Access to a Common Lexical Store. *Journal of Verbal Learning and Verbal Behavior, 22*(5), 526-534.

Potter, M. C., So, K. F., Von Eckardt, B., & Feldman, L. B. (1984). Lexical and Conceptual Representation in Beginning and Proficient Bilinguals. *Journal of Verbal Learning and Verbal Behavior, 23*(1), 23-38.

Scarborough, D. L., Gerard, L., & Cortese, C. (1984). Independence of Lexical Access in Bilingual Word Recognition. *Journal of Verbal Learning and Verbal Behavior, 23*(1), 84-99.

Schoonbaert, S., Duyck, W., & Brysbaert, M. (2003). *Associative and Translation Priming Symmetry in Lexical Decision and Episodic Recognition*. Unpublished Manuscript.

Trésor de la Langue Française. (1971). Nancy, France : CNRS.

Van Hell, J. G., & Dijkstra, T. (2002). Foreign Language Knowledge Can Influence Native Language Performance in Exclusively Native Contexts. *Psychonomic Bulletin & Review, 9*(4), 780-789.

Van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic Neighborhood Effects in Bilingual Word Recognition. *Journal of Memory and Language, 39*(3), 458-483.

Van Orden, G. C. (1987). A Rows Is a Rose - Spelling, Sound, and Reading. *Memory & Cognition, 15*(3), 181-198.

Van Orden, G. C., & Goldinger, S. D. (1994). Interdependence of Form and Function in Cognitive Systems Explains Perception of Printed Words. *Journal of Experimental Psychology: Human Perception and Performance, 20*(6), 1269-1291.

Van Wijnendaele, I., & Brysbaert, M. (2002). Visual Word Recognition in Bilinguals: Phonological Priming from the Second to the First Language. *Journal of Experimental Psychology: Human Perception and Performance, 3*, 619-627.

Weinreich, U. (1953). *Languages In Contact.* New York: The Linguistic Circle Of New York.

APPENDIX

Stimuli Collected by Brysbaert, Van Dyck and Van de Poel (1999, Appendix)

French (L2) Target	Dutch (L1) Homophonic	Dutch (L1) Graphemic Control	Dutch (L1) Unrelated Control
APTE	Abt	alt	olm
BASE	Baas	baan	rook
BATTE	Bad	bak	pil
BOUC	Boek	boot	deel
BOULE	Boel	beul	haak
CANE	Kan	dan	mug
CLOQUE	Klok	slot	small
COULE	Koel	doel	daad
COURS	Koer	roer	fooi
CRANE	Kraan	graan	stoom
DIRE	Dier	diep	taak
DOSE	Doos	doen	haat
DURE	Duur	durf	pijn
HUILE	Wiel	zeil	boon
ILE	Iel	iep	gok
MARE	Maar	maal	veel
NEZ	Nee	nek	oud
OUI	Wie	jij	dag
PART	Paar	paal	hoog
PATTE	Pad	pak	fel
PIRE	Pier	piek	kolf
PLACE	Plas	pias	huur
POTE	Poot	poos	jurk
POULE	Poel	poen	gist
RAME	Raam	raad	punt
RAVE	Raaf	rank	tolk
ROUTE	Roet	roes	haai
TOUT	Toe	tor	dag
VOUTE	Voet	volk	hard
ZONE	Zoon	zoen	kans

CHAPTER 6
WORDGEN: A TOOL FOR WORD SELECTION AND NON-WORD GENERATION IN DUTCH, GERMAN, ENGLISH, AND FRENCH

***Manuscript submitted for publication*[1]**

WordGen is an easy-to-use program that uses the CELEX and Lexique lexical databases for word selection and non-word generation in Dutch, German, English, and French. Items can be generated in these four languages, specifying any combination of seven linguistic constraints: number of letters, neighborhood size, frequency, summated position-nonspecific bigram frequency, minimum position-nonspecific bigram frequency, position-specific frequency of the initial and final bigram and orthographic relatedness. The program also has a module to calculate the respective values of these variables for items that have already been constructed (either with the program or taken from earlier studies). WordGen is especially useful for (1) Dutch and German item generation, because no such stimulus selection tool exists for these languages, (2) the generation of non-words for all four languages, because our program has some important advantages over previous non-word generation approaches and (3) psycholinguistic experiments on bilingualism, because the possibility of using the same tool for different languages increases the cross-linguistic comparability of the generated item lists.

---

[1] This paper was co-authored by Timothy Desmet, Lieven Verbeke and Marc Brysbaert.

INTRODUCTION

One of the most important stages in psycholinguistic research on word processing is the construction of items. To be able to draw valid and general conclusions on the basis of an experiment's outcome, the selection of words and non-words has to be performed with the utmost carefulness. Items have to be adequately manipulated on the experimental variables under scrutiny and items in different conditions have to be matched appropriately on potentially confounding factors. This paper presents WordGen, an easy-to-use tool that can substantially simplify and speed up the laborious job of item construction, which is mostly done manually up to now. WordGen uses the CELEX database (Baayen, Piepenbrock, & van Rijn, 1993, 1995) and the Lexique database (New, Pallier, Ferrand, & Matos, 2001) to generate word and non-word items in Dutch, German, English, and French. The program can be downloaded freely from the following URL: http://allserv.ugent.be/~wduyck/wordgen.htm. In order to install the program, the 1993 or 1995 CD-ROM version of the CELEX lexical database is needed. Upon installation, the CELEX lemma frequency databases of Dutch, English and German, are read from this CD-ROM, and parsed for use with WordGen. Because the Lexique database is freely available (http://www.lexique.org) and distributed under a GNU license, the data needed for French word and non-word generation are included in the program's download, so it is not necessary to download the Lexique database separately.

Before going into the details of the program and the underlying algorithms, we will briefly discuss the linguistic variables that can be controlled for by the program and their importance in the psycholinguistic literature. These variables include word frequency, neighborhood size, bigram frequency, orthographic relatedness, and word length.

One of the most important linguistic variables in word recognition is word frequency: words that occur more frequently are processed faster and more accurately than words that occur less frequently. This effect was first demonstrated in tachistoscopic recognition (Howes & Solomon, 1951) and was later generalized to a wide range of tasks, including lexical decision (e.g., Whaley, 1978) and word naming (e.g., Forster & Chambers, 1973). It is important to control for word frequency in psycholinguistic experiments because this variable has subtle effects, emerging not only between highly frequent and highly infrequent words, but even between frequent and slightly less frequent words. In the middle of the 1990s the suggestion has been made that all frequency effects in the literature were actually confounded age-of-acquisition effects (Morrison & Ellis, 1995). The age-of-acquisition of a word is the age at which a word is first learned (Carroll & White, 1973; Gilhooly, 1984). However, at present it seems that both frequency and age-of-acquisition have independent effects in word processing (e.g., Bonin, Chalard, Meot, & Fayol, 2001; Brysbaert, Lange, Van Wijnendaele, 2000; Gerhand & Barry, 1999; Morrison & Ellis, 2000). Word frequency is controlled or manipulated in virtually all word processing studies.

Another variable that affects word processing is orthographic neighborhood size. The neighborhood size of an item is the number of existing words that can be obtained by changing one letter of that item (Coltheart, Davelaar, Jonasson, & Besner, 1977). For instance, the Dutch word “klad” has six orthographic neighbors: “blad”, “glad”, “klak”, “klam”, “klap”, and “klas”. A large neighborhood size enhances the performance on naming and lexical decision, especially for low-frequency words (Andrews, 1989; Grainger, 1990; McCann & Besner, 1987). In non-word items, neighborhood size could be an indicator of how word-like a non-word is. For instance, an unpronounceable non-word such as “hzva” has no orthographic neighbors in English, whereas a slightly more pronounceable non-word such as “tuee“ has 5 neighbors, and a pseudoword such as “beed” has 14 neighbors.

Another lexical variable that our program allows to constrain is type bigram frequency. Bigrams are the adjacent letter pairs of an item. For instance, the word “code” has three bigrams: “co”, “od” and “de”. The effect of bigram frequency on word processing is a bit controversial. For instance, early effects of bigram frequency on word recognition (e.g., Broadbent & Rice, 1968; Orsowitz, 1963; Rice & Robinson, 1975; Rumelhart & Siple, 1974) were later argued to be confounded effects of subjective familiarity (e.g., Gernsbacher, 1984). Also, more recently some studies fail to find an effect of bigram frequency (e.g., Andrews, 1992) whereas others do find an effect (e.g., Westbury & Buchanan, 2001). Nevertheless, bigram frequency is still controlled for in numerous recent psycholinguistic studies (e.g., Bertram & Hyönä, 2003; Locker, Simpson, & Yates, 2003; Martensen, Maris, & Dijkstra, 2003; Yates, Locker, & Simpson, 2003). Moreover, from the perspective of this program, it is also an interesting variable to consider when making non-word items, because on average the higher the summated bigram frequency of a non-word, the more word-like it will be.

By allowing to indicate which letters should and should not be part of the generated items, WordGen also allows for the manipulation of the orthographic overlap between items. Numerous studies have found that orthographically related items (for instance “contrast” and “contract”) can prime each other (e.g. Brysbaert, 2001; Grainger & Ferrand, 1996; Van Heuven, Dijkstra, Grainger, & Schriefers, 2001). One of the earliest findings was that this orthographic priming effect only appears with masked primes (Humphreys, Besner, & Quinlan, 1988). More recently, Forster and Veres (1998) showed that the strength of this priming effect depends on the type of non-word distractors in the experiment. Interestingly, orthography not only influences visual word recognition processes, but has also been shown to play an important role in speech production (e.g., Damian & Bowers, 2003) and speech perception (e.g., Miller & Swick, 2003; Slowiaczek, Soltano, Wieting, & Bishop, 2003). Of course, with WordGen the orthographic overlap can not only be manipulated, but it can also be controlled for, which is of crucial importance in experiments that are exploring the independent effects of phonological or semantic priming.

Finally, our program also allows for the constraining of the length of a word or non-word by indicating the number of letters. Length effects have been demonstrated on lexical decision and naming times of words and non-words (e.g., Chumbley & Balota, 1984; Forster & Chambers, 1973; Weekes, 1997; Whaley, 1978). Virtually all word processing experiments control for length.

In the following sections, we will discuss how these variables have been implemented in WordGen. We will subsequently deal with the three different frames of the program: (a) WordGen options, (b) word / non-word generation and (c) word / non-word check.

WORDGEN OPTIONS

Before looking up word or non-word information in the "Check word / non-word" frame or generating items in the "word / non-word generation" frame, some options can be set. First, one of four languages needs to be selected: Dutch, German, English, or French. Next, WordGen allows for the output to be saved to a data file (the default file is called WWG_LOG.prn and is saved in the program's directory/folder). If this option is not chosen, the output only appears in the window on the right side of the program and is lost when the program is shut down. The user can also choose for the program to provide detailed output. When this option is selected, the extra information that is given consists of a list of all neighbors and the frequency of all separate bigrams of an item. Next, the non-word searching time can be limited to any number of seconds (and is set to 30 seconds as a default). This option is provided because asking the program for a non-word with a constraint combination which is too narrowly defined or even impossible can lead to an infinite (or very long-lasting) search. For instance, asking the program for a Dutch ten-letter non-word with fourteen neighbors and a very low summated bigram frequency is unlikely to be successfully completed within a reasonable time (if it is possible to find such a non-word at all). Thus, the program will keep on searching if the search time is not set to a specific limit. Users that are looking for non-words that have to meet certain strict – but not impossible - constraints should set the time limit very high or should deactivate it at all. In practice, the to-be-generated non-words will be matched to existing words, mostly leading to reasonable combinations of constraints. Note that this time limit does not apply to word generation because it does not take much time for the program to perform an exhaustive check of every word in the databases against the constraints that were set.

WORD / NON-WORD GENERATION

To generate a word the program registers the values of the constraints that were set by the user. The program randomly selects an entry in the CELEX or Lexique database and starts a serial search through the database looking for the first word that satisfies the combination of constraints provided by the user. Each time the user asks to generate a new word (even using the same parameter settings during the same session) a different random entry in the database is selected.

To generate a non-word, the program assembles a string of randomly selected letters and verifies whether the letter string is not an existing word in the lexical database for that respective language. Next, every constraint is checked and as soon as one of them is violated the random letter string is rejected and the process starts all over again until a letter string is assembled that conforms to all constraints or until the time limit that was set in the Options panel is reached. The latter case might be an indication that the parameters were set too narrowly and that the constraints should be broadened.

In practice, psycholinguists in the process of constructing non-words often base their non-words on existing words and change one letter to turn it into a non-word. This heuristic ensures that the non-words are mostly reasonably word-like. The program can be set to use this approach (see below), but we included the other (random letter generation) strategy as well because we believe it is desirable to allow for as much variation as possible in the type of non-words. For instance, we did not want to exclude possible non-words that had no neighbors but are still pronounceable word-like letter strings (e.g., “syspor”), a possibility that is excluded when basing non-words on existing words. Of course, we provided some other search options (see below) to make sure that the wordlikeness of the non-words generated by the program is preserved.

When generating a word or non-word, seven constraints can be set. The first and most straightforward constraint is the number of letters the generated item should have.

The second constraint is neighborhood size, or the number of orthographic neighbors an item can have. If this option is set, the program checks which words appearing in the respective CELEX/Lexique database have all but one letters in common with the candidate word/non-word. In this way a highly accurate count of the neighborhood size for a word in a given language is obtained. This is especially useful for Dutch and German, for which no neighborhood size norms are available at present. Hence, the program allows to avoid more elaborate and less accurate assessment strategies of neighborhood size which are often used in studies in these languages, such as asking a number of independent participants to name as many neighbors as possible of the items that will be used in the experiment (e.g., Van Hell & Dijkstra, 2002). When setting the neighborhood size constraint during a word/non-word search, it is important to know that neighborhood size is related to word length. For instance, whereas almost all three-letter words have at least a couple of neighbors, longer words mostly have zero, one or two neighbors. Appendix A shows the distribution of neighborhood size as a function of word length for Dutch, German, English and French.

This information might be useful when setting the neighborhood constraint. For example, looking at the first Figure of Appendix A, it is clear that it would not make much sense to ask the program for a Dutch eight-letter word with 12 neighbors. It should be noted that in the figures we left out the number of words with zero or one neighbors. Including this information would distort the scale of the Y-axis too much, because a huge amount of words have less than two neighbors. Note that this information would not be very useful anyway given the aim of these histograms, as searching for an item with less than two neighbors is never an unreasonable constraint.

The third constraint that can be set is the word frequency of an item. Obviously, this constraint can only be set in word generation. In our program the frequency of words is based on the lemma frequencies provided in the CELEX database for Dutch, German, and English, and the lemma frequencies provided in the Lexique database for French. In order to ensure a high comparability between different languages and possibly between different studies the program provides the logarithmic frequency per million words in the corpus. We decided to use lemma - and not for example wordform - frequencies for a number of reasons. First and most importantly, the former is by far most often used in psycholinguistic research. Second, the lemma database is smaller, which substantially speeds up the search process in the database (especially important for non-word generation, see further). Third, due to extensive manual coding and disambiguation the lemma database is more transparent with respect to its records than the wordform database. For example, the wordform databases contain a lot of compound records such as “go back on”. With respect to neighborhood size and bigram frequency calculation for instance, it is not desirable to take parts (e.g. “back”) of this record into consideration which are also dealt with as separate records. Any program resolving these issues is basically repeating part of the lemma coding. Fourth, due to its considerable size, it is very likely that the variables of interest to WordGen calculated on the basis of the lemma database, would correlate substantially with those based on the wordform database. Finally, several studies (e.g. Baayen, Dijkstra, & Schreuder, 1997; New, Brysbaert, Segui, Ferrand, & Rastle, 2003) suggest that the processing of words is partly driven by the frequency of morphologically related wordforms (e.g. plurals), which favors the lemma frequency approach (because these wordforms are grouped in the same lemma entry).

The fourth constraint is summated type bigram frequency. Our program summates the position-nonspecific frequency of each bigram of a letter string, based on how many times a bigram occurs in the CELEX or Lexique database independent of its position in the word. For example, the Dutch word “boek” has a bigram frequency of 19898, which is the sum of the number of occurrences of each of its bigrams: “bo” (4123), “oe” (9120), and “ek” (6655) in the CELEX. Because the four languages have a different number of words in the database there is a big difference between the bigram frequencies for these languages. For instance, the Dutch and English databases in the CELEX contain 124.136 and 52.447 entries, respectively. This means that on average Dutch summated bigram frequencies will be more than twice as high than English summated bigram frequencies. Also, because the program works with summated bigram frequencies, on average the bigram frequency for short words will be lower than the bigram frequency for long words. To help the user set the summated bigram frequency constraint we included four figures with the distribution information of bigram frequencies as a function of word length, plotted separately for Dutch, German, English and French (all Figures, see Appendix B).

Again, these figures should make clear that it does not make much sense, for instance, to ask for a Dutch four-letter word with a summated bigram frequency of at least 80.000. As a help for the user, the program adapts the default constraints for summated bigram frequency as a function of the language and the number of letters that was chosen. However, depending on the needs of the user it is advisable to look at the histograms in Appendix B to set this constraint to a more narrow range.

In addition to the summated bigram frequency constraints, the minimum ‘legal’ bigram frequency and the minimum position specific onset / suffix bigram frequency can also be set. These two constraints were added to increase the efficiency of constructing plausible non-words. If only the summated bigram frequency were constrained, it is possible that the program generates a non-word of which only one of its bigrams is highly frequent (leading to a high summated bigram frequency), but of which the other bigrams are totally infrequent in a given language, leading to an unpronounceable non-word. The minimum legal bigram frequency allows to indicate what the minimal bigram frequency should be for any of the bigrams of an item, so the user can make sure the non-word does not contain any infrequent bigrams which do not appear in any word in the respective lexical database. In practice, the default values set in the program have shown to be adequate. In addition, the onset and suffix position-specific bigram frequency can be constrained. This is because a lot of randomly generated non-words appear unusable because bigrams that are very frequent in other places in the word can still be very infrequent as the first or last two letters of a word. For instance, the bigram “rt” is very frequent in Dutch , but it is never the first bigram of a word. The position specific onset / suffix bigram frequency constraint makes sure that both the onset bigram and the suffix bigram occur a certain number of times as the onset or the suffix of a word. Hence, while the program includes the possibility to generate non-pronounceable non-words, it is strongly advised to search for parameter settings of these constraints which are adequate to the stimuli at hand, in order to obtain pronounceable non-words.

Additional to these bigram frequency constraints, we also included the possibility to use a widely adopted heuristic to enhance non-word generation even further (especially suited for non-words longer than seven letters).

When using the heuristic the program randomly selects an existing word and one letter is changed at a random position to turn it into a non-word that conforms to the other constraints that were set. This leads to very word-like non-words. Together with the bigram frequency constraints this heuristic ensures that non-words can be generated that vary widely between very word-like non-words to completely unpronounceable non-words.

As an illustration of how the different constraints settings influence the nature of the generated non-words, we ran a series of tests with different parameter settings. We generated four-, five-, six-, seven-, and eight-letter non-words either (1) with no constraints set at all, (2) with the minimum legal bigram frequency set at 30 and the minimum position-specific bigram frequency set at 15, (3) with the latter two constraints and , and the number of neighbors set to 1, or (4) using the heuristic without any other constraints set. For each of these conditions we let the program generate 100 Dutch non-words (for each of the different numbers of letters) and counted how many were pronounceable.

The results of this test are presented in Table 1. It is obvious from this table that using the minimum legal bigram frequency and the position-specific bigram frequency greatly improves the quality of the non-words compared to when no constraints are set. This is especially true when it is also requested that the non-words should have one neighbor. With these constraints, the ratio of pronounceable non-words is about 80%, which is quite high given the fact that the underlying algorithm only uses orthographic information, and does not have an extensive set of complicated grapheme-phoneme conversion rules. Hence, probing the program two times for a non-word will almost always result in a pronounceable non-word satisfying a combination of several lexical constraints. We believe this is a considerable improvement compared to classical non-word generation, which is often done manually (and therefore much slower) or pseudo-automated, without a clearly defined set of lexical criteria used to generate these items. This underspecification of non-word characteristics often makes it very difficult to compare non-word items across studies. This is especially troubling given the fact that changing the nature of filler non-words can influence the processing of the word stimuli, which is the actual object of interest (e.g. see the study of Forster & Veres, 1998, mentioned above, in which it was shown that the wordlikeness of used non-word targets interacts with the orthographic masked priming effect).

# letters	**Constraints**			
	no constraints[a]	*bigram*[b]	*bigram + neigbor*[c]	*heuristic*[d]
4	19	73	80	68
5	8	55	84	62
6	6	49	77	60
7	12	43	80	66
8	7	40	X*	74

Table 1. Number of pronounceable Dutch non-words (out of 100) as a function of number of letters and constraint settings.

[a] None of the constraints were set to a specific value.
[b] The minimum legal bigram frequency was set to 30; the minimum legal position-specific bigram frequency was set to 15.
[c] The minimum legal bigram frequency was set to 30, the minimum legal position-specific bigram frequency was set to 15, and the number of neighbors was set to 1.
[d] Only the heuristic was used; no other constraints were set.
X* Because random non-word generation takes very long in this condition (there are about 2 x 10^{11} possible random 8-letter string combinations), we advise considering the heuristic approach for non-words longer than seven letters.

The next constraint is the possibility to use a wildcard. This option allows the user to indicate whether the item should contain a specific letter on a specific position. For instance, a search for a five-letter word with a “p” on the second letter position can be indicated by typing *p*** next to the “use wildcard” option; a search for a seven-letter word with an “a” on the third position and an “s” on the fifth position can be asked for by **a*s**.

The final option is the forbidden letter list, which offers the possibility to indicate which letters should not be part of the generated item. If multiple letters need to be excluded, the letters should be typed next to each other without blank spaces or commas. For instance, when a word is needed that should not contain the letters m and r, the user should type mr next to the "forbidden letter list" option.

CHECK WORD / NON-WORD

In addition to the generation of words and non-words, the program also allows to calculate the respective values of the variables mentioned above for already constructed lists of words or non-words, either created with WordGen itself, or as a control of the stimuli of earlier studies. When checking an item the program verifies whether the item is a word or a non-word by looking whether or not it can be found in the CELEX or Lexique database. When the checked item is a word, the log frequency per million, the number of neighbors, and the summated type bigram frequency are provided. The same is true when the checked item is a non-word apart from the fact that the log frequency is not provided. When the detailed output option is requested the same information is provided, in addition to a list of the actual neighbors, and the frequency of each separate bigram of the item.

CONTRIBUTIONS TO THE FIELD

In the psycholinguistic literature there are a number of tools and databases available for stimulus generation. However, this is especially true for English and French and less so for Dutch and German. In this section, we will start give a concise overview of the most frequently used tools that are available for each of the four languages and we will outline the extra contribution of our tool for each of these languages.

In English, there is the MRC Psycholinguistic Database (Coltheart, 1981), which contains a large number of lexical properties of words, such as number of syllables, word frequency, imageability, age-of-acquisition, part of speech, stress pattern, etc. (see http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm). For the construction of non-word items there is the ARC non-word database, which contains monosyllabic non-word items that conform to the English phonological rules (Rastle, Harrington, & Coltheart, 2002; see also http://www.maccs.mq.edu.au/~nwdb/). We believe that our tool is complementary to both the MRC and ARC. For instance, the MRC Psycholinguistic Database does not contain summated bigram frequencies nor does it provide neighborhood size ratings. The ARC, on the other hand, does not contain multisyllabic non-words, and it only contains pseudohomophones or very word-like non-words, so it does not allow for as much variety in non-words as WordGen does. Hence, we believe WordGen may be useful (a) to calculate neighborhood size and bigram frequency measures of English words and non-words and (b) for the construction of English multisyllabic or low word-like non-words

In French there is the freely available Lexique database (New et al., 2001), which contains a huge amount of French lexical information, and LEXOP, a computerized lexical database which provides measures of the relationship between phonology and orthography for French monosyllabic words (Peereman & Content, 1999). Again, we think that our tool is an interesting extension to the French situation. First, no tool is available for non-word generation in French. Second, the availability of a number of different types of bigram frequency makes our tool very interesting for French stimulus generation.

Finally, our program is especially suited to be used in Dutch and German psycholinguistic experiments, because unlike English and French; these languages lack publicly available databases similar to the ones mentioned above, which contain frequently used lexical measures such as neighborhood size, bigram frequencies, and functions such as non-word generation. For instance, there are no available norms of neighborhood size for Dutch, forcing researchers to resort to inaccurate methods of controlling for neighborhood size, such as asking participants to name as many neighbors of the items that will be used in the experiment (see earlier). Now, more accurate norms are available, which can also be searched for by multiple entry points. It is not only possible to check how many neighbors a letter string has. It is also possible to ask the program for a word or non-word that has a specific number of neighbors. This advantage also holds for English and French, for which norms exist for words, but where it is harder to easily find words that have a prespecified number of neighbors (especially in combination with other lexical constraints).

Next to the fact that WordGen increases the possibility to generate words and non-words in Dutch, German, English, and French, our specific tool has some other advantages. First, this program is ideally suited for stimulus generation in the fast growing research domain of bilingualism. The same program and norms can be used to construct items in different languages, increasing the comparability of the item lists over languages. This is especially true given that the combination of different lexical variables can be constrained at the same time. The program can also easily be applied to any new language, given a reliable lexical database.

A final advantage of this program is that it allows for a great variation in non-word items, ranging from highly recognizable non-words to pseudowords. The way in which non-words are created traditionally – by taking a word and changing one letter – does not easily allow for the manipulation of wordlikeness (although this heuristic is also available in WordGen). This variation in wordlikeness is possible in WordGen by the specific way in which the non-words are constructed (creating random letter strings), which does not artificially excludes non-words that have no neighbors and are very non-word-like. Moreover, the possibility to specify bigram frequency and number of neighbors is a big advantage for researchers interested in the influence of non-word characteristics on the performance in word recognition tasks (e.g. Forster & Veres, 1998).

REFERENCES

Andrews, S. (1989). Frequency and neighborhood effects on lexical access: Activation or search? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 802-814.

Andrews, S. (1992). Frequency and neighborhood effects on lexical access – Lexical similarity or orthographic redundancy. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 234-254.

Baayen, R. H., Dijkstra, T. & Schreuder, R. (1997). Singulars and plurals in Dutch : evidence for a parallel dual-route model. *Journal of Memory and Language,* 37, 94-117.

Baayen, R.H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical data base (CD-ROM):* Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Baayen, R.H., Piepenbrock, R., & van Rijn, H. (1995). *The CELEX lexical data base (CD-ROM Second Release):* Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Bertram, R., & Hyönä, J. (2003). The length of a complex word modifies the role of morphological structure: Evidence from eye movements when reading short and long Finnish compounds. *Journal of Memory and Language, 48,* 615-634.

Bonin, P., Chalard, M., Meot, A., & Fayol, M. (2001). Age-of-acquisition and word frequency in the lexical decision task: Further evidence from the French language. *Cahiers de Psychologie Cognitive – Current Psychology of Cognition, 20,* 401-443. PT J

Brysbaert, M. (2001). Prelexical Phonological Coding of Visual Words in Dutch: Automatic After All. *Memory & Cognition, 29*(5), 765-773.

Brysbaert, M., Lange, M., & Van Wijnendaele, I. (2000). The effects of age-of-acquisition and frequency-of-occurrence in visual word recognition: Further evidence from the Dutch language. *European Journal of Cognitive Psychology, 12,* 65-85.

Carroll, J.B., & White, M.N. (1973). Age of acquisition norms for 220 picturable nouns. *Journal of Verbal Learning and Verbal Behavior, 12,* 563-576.

Chumbley, J.I., & Balota, D.A. (1984). A words meaning affects the decision in lexical decision. *Memory and Cognition, 12,* 590-606.

Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology, 33A,* 497-505.

Coltheart, M., Davelaar, E., Jonasson, J.T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and Performance VI* (pp. 535-555). London: Academic Press.

Damian, M.F., & Bowers, J.S. (2003). Effects of orthography on speech production in a form-preparation paradigm. *Journal of Memory and Language, 49,* 119-132.

Forster, K.I., & Chambers, S.M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior, 12,* 627-635.

Forster, K.I., & Veres, C. (1998). The prime lexicality effect: Form-priming as a function of prime-awareness, lexical status, and discrimination difficulty. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 498-514.

Gerhand, S., & Barry, C. (1998). Word frequency effects in oral reading are not merely age-of-acquisition effects in disguise. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 267-283.

Gernsbacher, M.A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology – General, 113,* 256-281.

Gilhooly, K.J. (1984). Word age-of-acquisition and residence time in lexical memory as factors in word naming. *Current Psychological Research, 3,* 24-31.

Grainger, J. (1990). Word-frequency and neighborhood frequency-effects in lexical decision and naming. *Journal of Memory and Language, 29,* 228-244.

Grainger, J., & Ferrand, L. (1996). Masked Orthographic and Phonological Priming in Visual Word Recognition and Naming: Cross-Task Comparisons. *Journal of Memory and Language, 35*(5), 623-647.

Howes, D.H., & Solomon, R.L. (1951). Visual duration threshold as a function of word probability. *Journal of Experimental Psychology, 41,* 401-410.

Humphreys, G.W., Besner, D., & Quinlan, P.T. (1988). Event perception and the word repetition effect. *Journal of Experimental Psychology: General, 117,* 51-67.

Locker, L., Simpson, G.B., & Yates, M. (2003). Semantic neighborhood effects on the recognition of ambiguous words. *Memory and Cognition, 31,* 505-515.

Martensen, H., Maris, E., & Dijkstra, T. (2003). Phonological ambiguity and context sensitivity: On sublexical clustering in visual word recognition. *Journal of Memory and Language, 49,* 375-395.

McCann, R.S., & Besner, D. (1987). Reading pseudohomophones – implications for models of pronunciation assembly and the locus of word frequency effects in naming. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 14-24.

Miller, K., & Swick, M.D. (2003). Orthography influences the perception of speech in alexic patients. *Journal of Cognitive Neuroscience, 15,* 981-990.

Morrison, C.M., & Ellis, A.W. (1995). Roles of word-frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology – Learning, Memory, and Cognition, 21,* 116-133.

New, B., Brysbaert, M., Segui, J., Ferrand, L., & Rastle, K. (2003). *The Processing of singular and plural nouns in French and English*. Manuscript submitted for publication.

New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). A lexical database for contemporary french on internet: Lexique. *Année Psychologique, 101,* 447-462.

Peereman, R., & Content, A. (1999). LEXOP: A lexical database providing orthography-phonology statistics for French monosyllabic words. *Behavior Research Methods, Instruments, & Computers, 31,* 376-379.

Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC Nonword Database. *Quarterly Journal of Experimental Psychology, 55A,* 1339-1362.

Rice, G.A., & Robinson, D.O. (1975). Role of bigram frequency in perception of words and nonwords. *Memory and Cognition, 3,* 513-518.

Rumelhart, D.E., & Siple, P. (1974). Process of recognizing tachistoscopically presented words. *Psychological Review, 81,* 99-118.

Slowiaczek, L.M., Soltano, E.G., Wieting, S.J., & Bishop, K.L. (2003). An investigation of phonology and orthography in spoken-word recognition. *Quarterly Journal of Experimental Psychology, 56A,* 233-262.

van Hell, J.G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review, 9,* 780-789.

Van Heuven, W. J. B., Dijkstra, T., Grainger, J., & Schriefers, H. (2001). Shared Neighborhood Effects in Masked Orthographic Priming. *Psychonomic Bulletin & Review, 8*(1), 96-101.

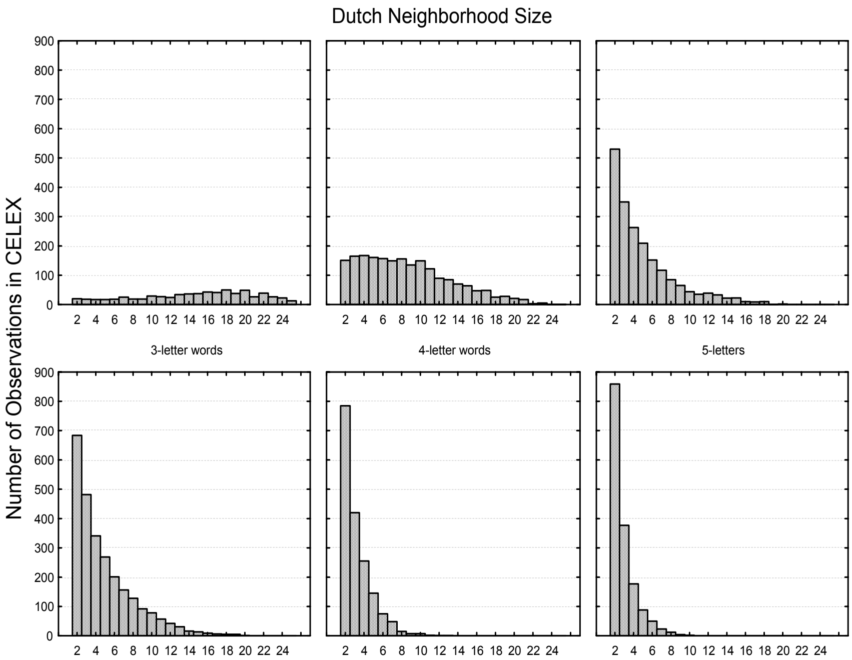
Weekes, B.S. (1997) Differential effects of number of letters on word and nonword naming latency. *Quarterly Journal of Experimental Psychology, 50A,* 439-456.

Westbury, C., & Buchanan, L. (2002). The probability of the least likely non-length-controlled bigram affects lexical decision reaction times. *Brain and Language, 81,* 66-78.

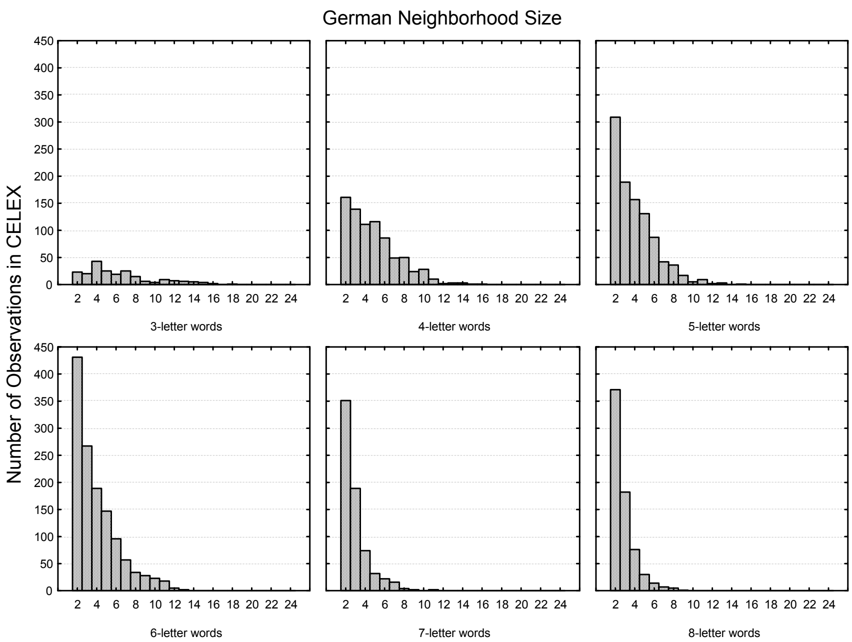
Whaley, C.P. (1978). Word - nonword classification time. *Journal of Verbal Learning and Verbal Behavior, 17,* 143-154.

Yates, M., Locker, L., & Simpson, G.B. (2003). Semantic and phonological influences on the processing of words and pseudohomophones. *Memory and Cognition, 31,* 856-866.

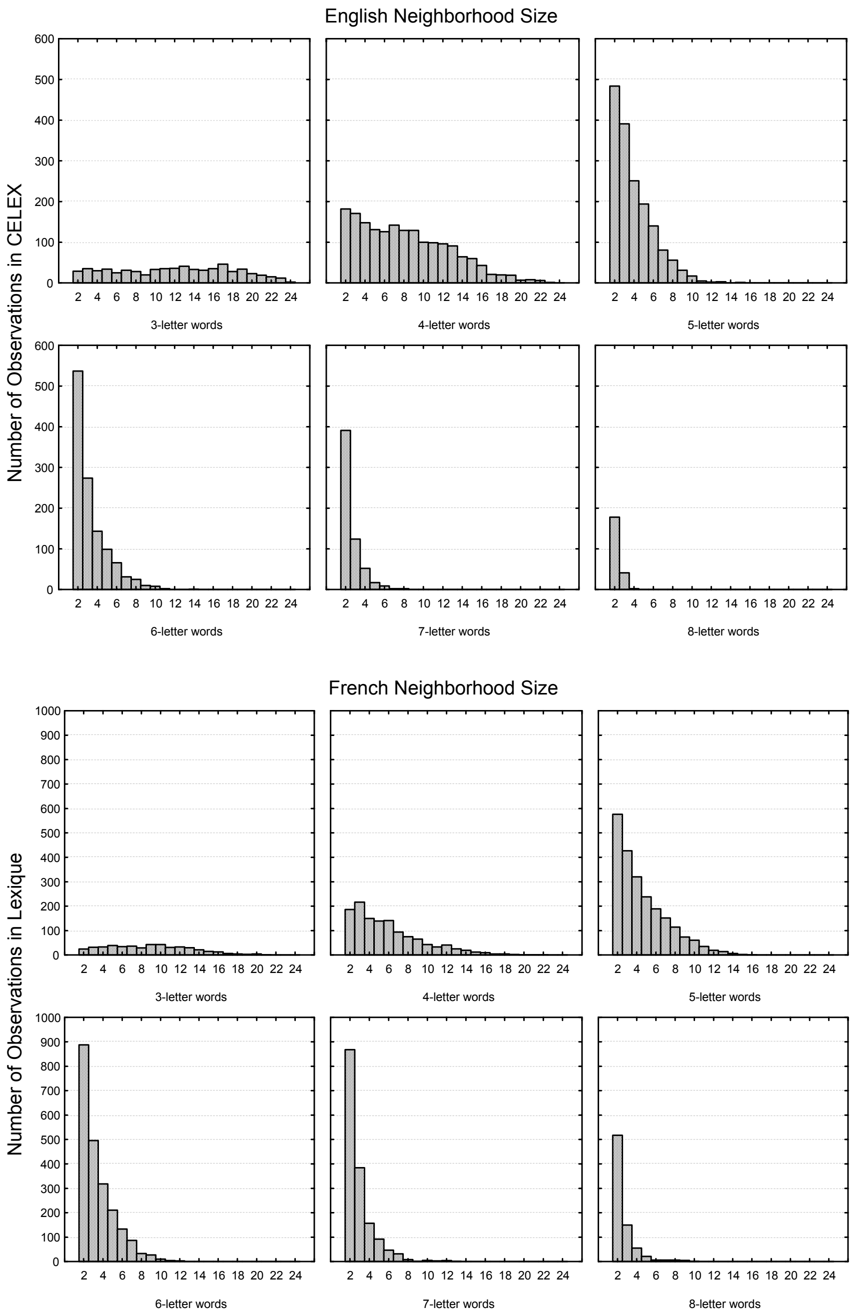
APPENDIX A



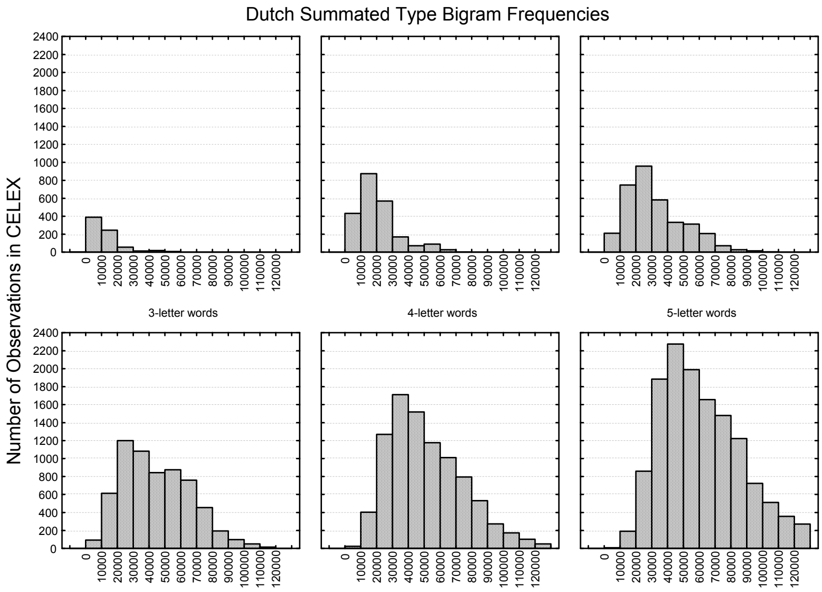
Dutch Neighborhood Size
Number of Observations in CELEX
3-letter words
4-letter words
5-letters

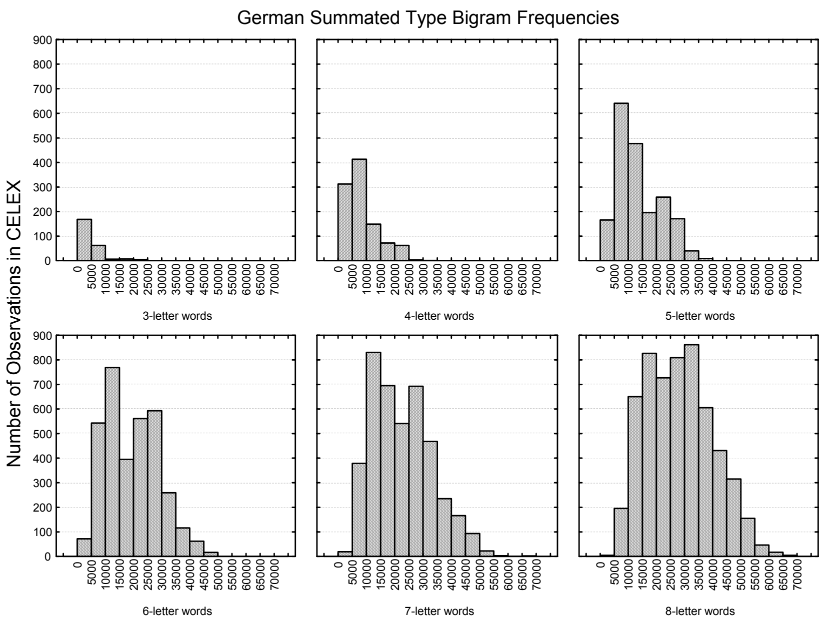


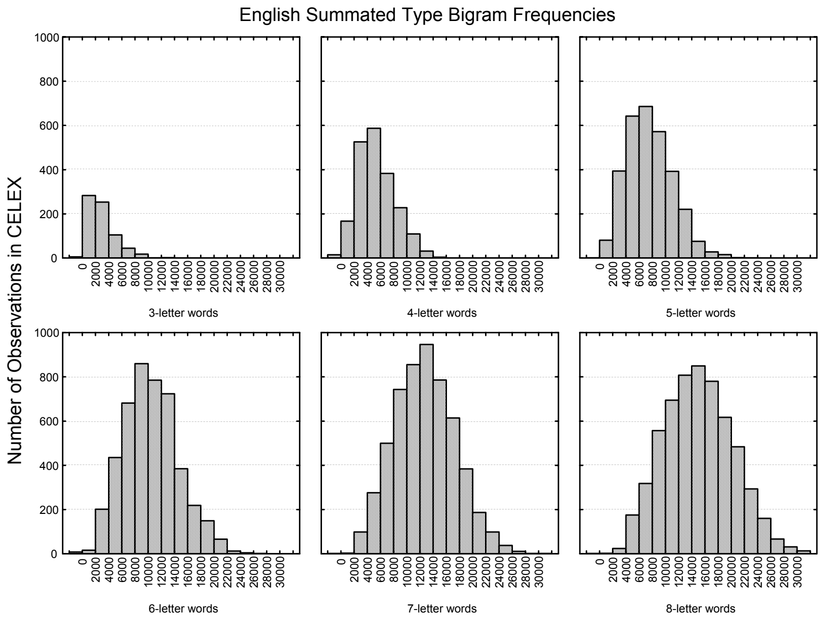
German Neighborhood Size
Number of Observations in CELEX
3-letter words
4-letter words
5-letter words
6-letter words
7-letter words
8-letter words

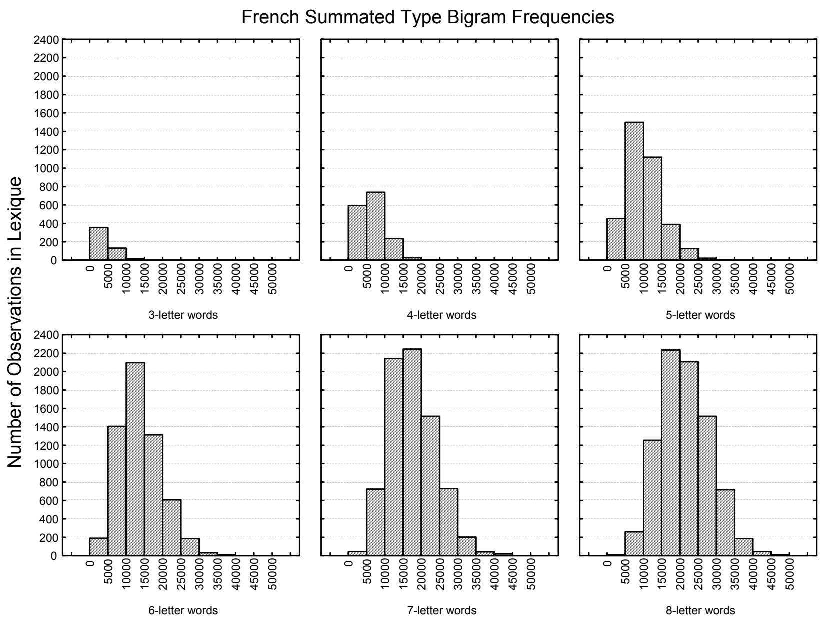

English Neighborhood Size
Number of Observations in CELEX
0
100
200
300
400
500
600
2 4 6 8 10 12 14 16 18 20 22 24
3-letter words
4-letter words
5-letter words
6-letter words
7-letter words
8-letter words
French Neighborhood Size
Number of Observations in Lexique
0
100
200
300
400
500
600
700
800
900
1000

APPENDIX B


Dutch Summated Type Bigram Frequencies
Number of Observations in CELEX
3-letter words
4-letter words
5-letter words


German Summated Type Bigram Frequencies
Number of Observations in CELEX
3-letter words
4-letter words
5-letter words
6-letter words
7-letter words
8-letter words


English Summated Type Bigram Frequencies
Number of Observations in CELEX
3-letter words
4-letter words
5-letter words
6-letter words
7-letter words
8-letter words


French Summated Type Bigram Frequencies
Number of Observations in Lexique
3-letter words
4-letter words
5-letter words
6-letter words
7-letter words
8-letter words

CHAPTER 7
GENERAL CONCLUSIONS

The aim of the experimental research presented in this doctoral dissertation was threefold. First, we tried to gain further knowledge with respect to the strength of L1 and L2 lexico-semantic connections. Secondly, it was investigated, using a selective interference paradigm, whether other than lexical associations are formed during the first stages of word acquisition. Thirdly, we examined whether pre-lexical access to phonological representations in bilingual visual word recognition is language-independent. In this final chapter, the main empirical findings of this thesis are summarized and the implications for current and future models of bilingualism are discussed. The chapter is concluded with some directions for future experimental investigations of bilingual language representation.

RESEARCH OVERVIEW

LEXICO-SEMANTIC CONNECTIONS: STRENGTH AND DEVELOPMENT

In the first and second part of this dissertation (Chapter 2 and 3), we explored the strength and development of L1 and L2 lexico-semantic connections. The dominant view on this issue is provided by the RHM of Kroll and colleagues (e.g. Kroll & Stewart, 1994; Kroll & de Groot, 1997), described in the introductory chapter of this dissertation.

According to the RHM, the links between L1 word forms and the meaning they represent are stronger than those between L2 word forms and their semantic representations. Similarly, lexical word-word connections are thought to be stronger from L2 to L1 than in the opposite direction (the asymmetry hypothesis). Hence, L2 word forms are assumed to access the semantic system primarily through associations with their translation equivalents at the lexical level. Consequently, the RHM predicts that forward translation is more likely to engage semantic activation than backward translation, certainly at low levels of L2 proficiency. Also, backward translation is predicted to be faster, because translation equivalents are accessed directly, without intermediate semantic activation. This asymmetry in the strength of lexico-semantic connections is believed to weaken in high levels of L2 proficiency (the developmental hypothesis). Whereas the RHM has done a very good job in explaining a wide range of empirical results, the experiments presented in Chapter 2 and 3 suggest that an update of the model may be warranted, as we obtained evidence against the model's asymmetry and developmental hypotheses.

STRENGTH OF LEXICO-SEMANTIC CONNECTIONS

All four experiments of Chapter 2 showed clear semantic effects of number magnitude in forward, but also in backward translation. Thus, for Dutch-French bilinguals, it took less time to forward translate *twee* (L1) into *deux* (L2) [two] than *acht* (L1) into *huit* (L2) [eight]. Similarly, they were faster to backward translate *deux* into *twee* than *huit* into *acht*. Because information with respect to the primary meaning (i.e. quantity) that a (L2 or L1) number represents is not stored at the lexical level, this strongly indicates semantic activation during the translation process. Moreover, this magnitude effect was equally strong in both translation directions, and was not present when the number words had to be named instead of translated. This is clear evidence that the mappings between L2 number word forms and their underlying semantic representations (a) exist and (b) are activated fast en strongly enough to influence the translation process. These findings are not compatible with the asymmetry in the RHM, according to which lexico-lexical connections from L2 to L1 should significantly be stronger than L2 lexico-semantic connections.

These findings are not without analogues in the literature. Similar findings for number word stimuli were recently obtained by Duyck and Brysbaert (2002). Using a masked priming paradigm, they showed that the translation of L1 (e.g. *vijf*) and L2 (e.g. *cinq*) target number words is facilitated when they are preceded by an Arabic digit prime from which the magnitude it represents is close to that of the target (e.g. *4*). Priming of a more distant Arabic digit (e.g. *2* or *3*) was smaller, but still significant. These semantic priming effects were equally large for both directions of translation. Similar indications of strong L2 lexico-semantic connections for other types of L2 words were also recently obtained by several authors (Bloem & La Heij, 2003; Francis, Augustini, & Saenz, 2003; La Heij, Hooglander, Kerling, & Van der Velden, 1996). La Heij et al., for instance, found that both forward and backward translation of target words was facilitated by distractor pictures of an object (e.g. a chair) belonging to the same semantic category as the target (e.g. *table*).

DEVELOPMENT OF LEXICO-SEMANTIC CONNECTIONS

We also obtained evidence against the developmental hypothesis of the RHM, which states that the strength of the L2 lexico-semantic connections only notably increases in high L2 proficiency levels.

First, in Chapter 2, we replicated the semantic number magnitude effects discussed above with participants who learned a set of 'Estonian' number words only a few minutes before the experiments. This suggests that new L2 (number) word forms are mapped onto semantic information much earlier in the L2 acquisition process than assumed by the RHM. This was the case even though learning occurred through associative word learning, which pre-eminently tempts to form lexical word associations.

Secondly, the developmental hypothesis of the RHM was also refuted in Chapter 3, using a totally different experimental paradigm. Duyck et al. (2003) started from the common finding in working memory literature that verbal working memory is not involved in associative learning of known word – word pairs, whereas the opposite is true for known word – new word pairs. This suggests that new words are learned (at least in associative word learning) by learning associations (through verbal working memory) between the phonological representations of both word forms (as assumed by the RHM[1]). Duyck et al. noted that the fact that these phonological associations were apparently not formed when learning associations between known words, may be due to the fact that the associations are learned

[1] Note that the RHM assumes that L2 words are initially learned through lexical associations with their L1 translation equivalents, but is neutral with respect to which lexical representations (i.e. orthographic or phonological) are primarily involved in this process.

through non-lexical (visual) associations. This was confirmed in Experiment 1 (see Chapter 3), in which we showed, using a selective interference paradigm, that such form-level associations were indeed formed if one of the known words had a low imageability. More importantly for the issue at hand, Experiment 2 showed that associations between a known word (a 'translation equivalent') and a new word are learned through visual (non-lexical) associations if a visual representation is available. Hence, this suggests that new word forms are mapped onto existing semantic representations early during word acquisition. Again, similar to the findings obtained in the 'Estonian' experiment described above, this is not compatible with the developmental hypothesis of the RHM. Also, it shows that the early lexico-semantic mapping observed in Chapter 1, is probably not restricted to words from which the meaning is so clearly defined and overlapping across languages as is the case for number words.

Moreover, the semantic information in this last study was of a visual nature, which is compatible with the work of de Groot and colleagues. In the distributed feature model (e.g. de Groot, 1992; de Groot, Dannenburg, & van Hell, 1994b; van Hell & de Groot, 1998c), it is assumed that (L1 and L2) words are represented semantically through a number of feature nodes. The degree of overlap in meaning between translation equivalents, indexed by the number of shared features, depends on the type of word being represented. One of the main tenets of the distributed feature model is that concrete words (e.g. *ball*) have more similar meanings (indicated by a larger feature overlap) across languages than abstract words (e.g. *justice*), and thus have more (stronger) lexico-semantic connections that can activate their translation equivalents through the semantic system. The findings of Chapter 3 suggest that these shared semantic representations on which L2 word forms are mapped are indeed of a visual nature.

Again, the indications of early form-to-meaning mappings obtained in Chapter 2 and 3, although incompatible with the RHM, are not without parallels. A few studies found similar indications of early L2 lexico-semantic mappings, both for numbers and other types of words. As for number words, Tzelgov, Yehene, Kotler, and Alon (2000) showed that new nonsense symbols which are taught to represent magnitudes from 1 to 9 exhibit clear semantic number magnitude effects (i.e. the distance effect: it is easier to make a magnitude comparison between numbers which are further apart) very fast.

As for other types of words, similar findings were reported by Altarriba and Mathis (1997). After training a group of monolinguals on a set of English – Spanish word pairs, they found more errors on lexically, but also on semantically related false translations than on unrelated words in a translation recognition task. Again, this suggests that the newly learned Spanish word forms could already activate semantic representations to some degree, causing the interference effect. In a more recent study by von Pein and Altarriba (2003), similar findings were reported for English participants learning non-iconic American Sign Language gestures.

For further theoretical implications of Chapter 2 and 3, I refer to the *Theoretical Implications* section later in this chapter.

NON-SELECTIVE ACCESS TO PHONOLOGICAL REPRESENTATIONS

In the third part of this dissertation (Chapter 4 and 5), we explored whether access to phonological representations in bilingual visual word recognition is language selective. As already stated in the introductory part of this chapter, the few studies which have tackled this issue suggest that access to phonological representations is language-independent, just like access to orthographic lexical representations (e.g. the BIA model, for a review see Dijkstra & Van Heuven, 2002). Our findings offer further evidence for a strong language non-selective view on phonological coding in bilinguals. In the present section, we will respectively deal with L1 influences on L2 processing and vice versa.

L1 PHONOLOGICAL ACCESS DURING L2 WORD RECOGNITION

First, in Chapter 4, we showed that the recognition of a L2 (English) word (e.g. BACK) is facilitated by a L1 (Dutch) masked prime (e.g. *ruch*) which is a pseudohomophone of its translation equivalent (e.g. rug). This finding is in fact very similar to a study of Tan and Perfetti (1997; but see Zhou & Marslen-Wilson, 1999), who demonstrated that a Chinese target word can be primed with a homophone of a target synonym. If one starts from a strong non-selective view on lexical access in bilinguals (see earlier), there is not much difference between an 'intra-lingual' synonym and a 'cross-language' translation equivalent, in that they are both different lexical labels representing the same meaning.

Secondly, we replicated the above effect with L2 targets (e.g. CHURCH) and L1 pseudohomophone associative primes (e.g. *pous* [paus – pope]) (instead of pseudohomophone translation primes).

Thirdly, we also found evidence of L1 influences on L2 processing in the absence of L1 stimuli (although the stimuli in the previous experiments were masked primes). More specifically, the recognition of L2 words (e.g. CORNER) was facilitated by more frequent L2 homophones (e.g. *hook*) of their L1 translation equivalents (e.g. hoek). This suggests that ambiguous L2 phonological representations (interlingual homophones, e.g. /huk/) quickly activate all underlying semantic representations, even if they correspond to two different languages and are not related (e.g. [hook]-[corner]). In this case for instance, the phonological representation /huk/ activated its L1 meaning [corner], even though the experiment only contained L2 stimuli.

Because the words and nonwords in these three experiments were only homophones of the translation equivalent according to GPC rules of the non-target language, this strongly suggests that L1 GPC rules[2] are active during L2 word recognition. Also, the pseudohomophone associative priming effect shows that the overlap between the concepts that the prime and the target represent needs not to be complete (as is the case for translation equivalents) in order for the priming effect to arise. Hence, the activation in L1 phonological representations is large enough to spread to related concepts (see also the *Implications for Models of Bilingualism* section).

Finally, in Chapter 5 (Duyck, Diependaele, Drieghe, & Brysbaert, 2004), we investigated whether the (Dutch-French) cross-lingual pseudohomophone priming effect (i.e. recognition of the L2 target pour is facilitated by the L1 pseudohomophone poer, see Brysbaert, Van Dyck, & Van de Poel, 1999) interacts with L2 proficiency. This was motivated by a claim of Gollan, Forster and Frost (1997), who hypothesized that reliance on phonology is negatively correlated with L2 proficiency. If this is true, one should expect a smaller phonological priming effect in highly proficient, balanced bilinguals than in late bilinguals. However, we found no evidence for this hypothesis. This shows that the importance of phonological coding during L2 word recognition is not restricted to the early stages of L2 acquisition.

These L1 influences are fully compatible with much of the earlier work on phonological coding in bilinguals (see Chapter 1). For instance, Jared and Kroll (2001) found that L2 words which have word-body enemies in L1 (e.g. the English word *bait* contains the word body *ait* which is pronounced differently in French) were named slower than control words. Also, Dijkstra et al. (1999) reported that phonological overlap of L2 target words with L1

---

[2] As already stated in Chapter 1, I would like to note that the term 'GPC rules' does not imply that grapheme-to-phoneme conversion follows strict, mutually exclusive 'rules', according to which one 'rule' eventually activates one phonological representation. Rather, it is believed that phonological conversions occur in parallel (across languages).

words (not shown in the experiment) inhibited recognition of those target words, even though L1 phonology was not relevant for the task at hand.

L2 PHONOLOGICAL ACCESS DURING L1 WORD RECOGNITION

Although it might be very plausible that L1 phonological representations are always activated to some degree (and can not be 'turned off'), the opposite is true for L2. One does not expect an influence of L2 in a native language setting. However, in Chapter 4, we replicated the pseudohomophone translation priming effect for L1 targets (e.g. TOUW) and L2 pseudohomophone translation primes (e.g. *roap* [rope]). Moreover, the prime effect was equally large from L2 to L1 as from L1 to L2. However, in contrast with L2, the pseudohomophone associative priming effect was not significant with L1 targets (e.g. BEEN [leg]) and L2 primes (e.g. *knea*). Also, L1 targets (e.g. *DAG*) were not processed faster if they were preceded by intra-lingual homophones (e.g. *dij*) of their L2 translations (day). Given the large body of evidence for pre-lexical phonological coding in L1 word recognition (e.g. Frost, 1998), the absence of an effect here probably is due to the fact that the mapping from an ambiguous phonological code (e.g. /dei/) on its L2 meaning [day] is much weaker than the mapping from phonology on the L1 meaning [thigh] (*dij*).

The first finding is in line with recent work from Van Wijnendaele and Brysbaert (2002), who showed that the recognition of L1 words (e.g. OUI [yes]) in French-Dutch bilinguals is facilitated by L2 homophone primes (e.g. *wie* [who]). This also strongly suggests that words access their L2 phonological representations, even when performing a task in the native language. Consequently, it can be assumed that phonological coding is fundamentally language-independent. However, the weaker (or absent) effects found in the other two experiments discussed above (e.g. pseudohomophone associative priming) suggest that the overall (or rest) activation in L2 lexical representations is lower than for L1. This is compatible with Jared and Kroll (2001, see Chapter 1), who found that L1 target words having word body enemies (see earlier) in L2 were only named slower if participants had just named a block of L2 filler words. Further theoretical implications of Chapter 4 and 5 will be discussed in the next section.

IMPLICATIONS FOR MODELS OF BILINGUALISM

The empirical findings obtained in the present dissertation have several implications for present and future models of bilingualism. As a starting point of the present section, I will discuss which amendments could be made to the RHM (Kroll & Stewart, 1994) in order to cope with the current observations on L2 lexico-semantic organization. Then, it will be discussed how the model could be extended to include phonological and orthographic lexical representations. Of course, the model sketched below is still a hypothetical description, from which several assumptions still need additional empirical evidence.

First, as already stated in Chapter 2, the RHM is implicitly based on the horserace metaphor. Translation either follows the lexical route (in backward translation) or the semantic route (in forward translation). The fastest route completely determines the output. If this were replaced by a mechanism of parallel activation in a connectionist-type model, the central question should be *how much* each of the routes contributes to the build-up of the overall output activation. In this view, one route is not faster than the other; it only may have stronger connection weights and, therefore, influence the activation of the output units to a larger degree. Within this framework, the connection weights can be of such a nature that the functioning of such a model is very close to that of the classical RHM.

A second proposal to improve the current theoretical framework of the RHM concerns the fact that the asymmetry now only depends on L2 proficiency.

Because of this, the model predicts the same semantic involvement for the translation of all types of words, including number words, abstract words, and even syntactic function words. For this reason too, our findings with number words (Chapter 2) are critical for the model as a whole (Tzelgov, Yehene, Kotler, & Alon, 2000). It seems to us that the RHM should provide room for influences of both word and participant variables. So, the connection weights between lexical and semantic representations could depend on the consistency of the word's meaning, and therefore be larger for words that always have the same meaning (e.g. *three*) than for words that have different meanings as a function of the context (e.g., *great*). In this view, the impact of the semantically mediated route on translation times would also depend on the degree of semantic overlap between two translation equivalents. In turn, this overlap could be depending on word variables. Note that this is in fact very similar to the distributed feature model of de Groot and colleagues (see earlier, e.g. de Groot, 1992), according to which semantic activation in translation depends on semantic overlap, partially determined (de Groot, 1993; de Groot, Dannenburg, & van Hell, 1994a; Van Hell & de Groot, 1998a; Van Hell & de Groot, 1998b) by word concreteness (with concrete translation equivalents sharing more features than abstract words). This line of reasoning might explain the strong semantic effects obtained in Chapter 2, as the stimuli used in studies supporting the asymmetry hypothesis probably have less similar meanings across languages than number words. Note that variables such as cross-lingual overlap and word concreteness could not only influence the strength of lexico-semantic connections, but also the speed of development. This could explain the early form-to-meaning mappings observed respectively in Chapter 2 (Experiment 3) and Chapter 3 (Experiment 2). Also, connection weights could not only differ as a function of word variables between the lexical and the semantic level, but possibly also between the L1 and L2 lexical representations. These connections are probably stronger for translation equivalents with a large form overlap (so-called cognates, e.g. *ball* - *bal* for an English-Dutch bilingual) than for words with a small form overlap (e.g. *duty - plicht*). This could explain the fact that words with a large form overlap (so-called cognates) are easier to translate and show less evidence for semantic mediation in the translation process, than words with no form overlap (e.g. de Groot, 1992). For a more detailed account of this hypothetical model, we refer to Chapter 2 (see also this chapter's Figure 7).

Thirdly, we believe any future model of bilingualism should contain phonological representations, in addition to orthographic and semantic ones. We already noted in the beginning of this dissertation that no model of bilingualism at present does. Still, there is room for phonology in the sketch of the recent BIA+ model (Dijkstra & Van Heuven, 2002, see Chapter 1), but this system has not been implemented yet. From the present dissertation (Chapter 4 and 5), it will be clear that this system will have to be fundamentally non-language selective and highly interacting with semantic and lexical representations (discussed above) in order to cope with the results of previously mentioned studies. For instance, in Chapter 5, it was shown that phonological codes are assembled and activate semantic (and lexical) representations from both languages very fast. At present, probably the most important assumption that the authors have made with respect to this subsystem is the temporal delay assumption. This states that L2 phonological (and semantic) representations are accessed slower than L1 codes. The present dissertation suggests that this assumption may be too strong (see also Brysbaert, Van Wijnendaele, & Duyck, 2002) and that relative speed of phonological activation may be less language-dependent. For instance, the cross-lingual pseudohomophone translation priming effect obtained in Chapter 4 was not significantly stronger from L1 to L2 (Experiment 1) than from L2 to L1 (Experiment 2). Also, Experiment 5 showed a priming effect of L2 primes (homophones of the L2 target's translation equivalent), whereas no such effect was found in Experiment 6 (L1 primes). These observations suggest that L2 phonological processing is not always necessarily running behind on L1 phonological activation. This does not imply however, that connections from and to L2 representations are always as strong as for L1. In Chapter 5 for example, we showed that ambiguous phonological representations (from interlingual homophones) always activate their L1 meaning, even when performing a L2 task, whereas there was no sign of those phonological representations activating their L2 meaning when reading in L1.

Finally, as for the organization of the lexical (orthographic) representations in the model described above, I can be short. Given the large body of evidence supporting the BIA model (for a review, see Dijkstra & Van Heuven, 2002), this is probably the most powerful available account of this issue at present.

To conclude this section, I would like to point to some recently obtained empirical evidence (not included in the present dissertation), supporting an assumption made in the model above, with respect to the influence of word concreteness on the strength of lexico-semantic connections (Schoonbaert, Duyck, & Brysbaert, 2004). In this study, we showed that the inconsistency in the literature with respect to the existence of the backward translation priming effect (i.e. priming from a L2 prime on its L1 translation equivalent, e.g. see Grainger and Frenck-Mestre, 1998, versus Gollan et al., 1999) may be due to a confound of word concreteness in earlier studies. Using a lexical decision task with Dutch-English bilinguals, we found that the recognition of L1 concrete target words (e.g. SLEUTEL) is facilitated by (concrete) L2 translation primes (e.g. *key*). Such an effect was not found for abstract prime/target pairs (e.g. *honor* – EER). This indicates (a) that the translation priming effect is of a semantic nature[3] (see also the associative-semantic discussion in Chapter 4) and (b) that highly imageable words indeed have stronger lexico-semantic connections than abstract words.

---

[3] Because lexical representations only contain form-related information, a priming effect which is of a lexical/associative nature should not interact with a semantic variable such as word concreteness. See also the discussion about the semantic vs. lexical/associative nature of translation priming effects in the *General Discussion* section of Chapter 4.

To summarize, in the present section we have argued for a model of bilingualism (a) which is based on a parallel (connectionist-type) activation mechanism, (b) which allows influences of word-level variables on the developmental speed and strength of phonological, orthographic, semantic and interrepresentational connections and (c) in which access to the highly interacting representational (orthographic, semantic and phonological) subsystems is fundamentally non-language selective. To conclude this dissertation, some avenues for future experimental research on bilingualism will be discussed.

FUTURE RESEARCH

First, it would certainly be interesting to examine whether the currently dominant view on non-selective orthographic/lexical access in bilinguals, and the similar standpoint for phonological representations defended in this dissertation, can be generalized to sentence processing. In all studies on these issues, words were always presented isolated, separated from any meaningful linguistic context. This seriously limits the ecological validity of these findings, because bilinguals (as monolinguals) practically always encounter L2 (and L1) words within a contextually rich setting (e.g. a sentence), which may influence lexical access and recognition of the involved words. For instance, it is possible that recognition of the word *room* in the sentence '*now go upstairs to your room'* by a Dutch-English bilingual is not influenced by the fact that the word *room* also exists in Dutch [cream], in contrast to what would be observed in simple word recognition (e.g. Dijkstra et al., 1999). The same applies for phonological representations. It is possible that naming by a French-English bilingual of a L2 word like *bait* (containing the English word body *ait*, which is pronounced differently in French, L1) in the sentence '*the fish took the bait*' is not slower than naming of a control word. Still, this is exactly what Jared and Kroll (2001) found in a word recognition study (see earlier). Future studies will have to show whether the degree of semantic restrictedness of a sentence, or the predictability of a word in it, influences orthographic and phonological lexical access.

Secondly, we already noted that there is great inconsistency in the literature with respect to the existence of a backward translation priming effect. Grainger and Frenck-Mestre (1998) for example, found that the recognition of a L1 target word is facilitated by a L2 translation prime, whereas Gollan et al. (1997) found the opposite. This is an important issue because the latter finding is often used as an argument to defend the asymmetry hypothesis of the RHM, discussed earlier. It is then assumed that the L2 prime does not access its semantic representation due to weak lexico-semantic links, and therefore is unable to pre-activate the target. As noted in the previous section, we recently obtained strong evidence that the backward translation priming effect interacts with the word variable concreteness. This suggests an influence of this variable on the nature of the word's representation. Future studies will have to indicate whether there are other word variables which may also have an impact on L2 lexico-semantic organization, and therefore on the backward translation priming effect. This could be the case for example, for age of acquisition, frequency and familiarity.

Thirdly, future work should definitively identify the locus of such cross-language priming effects. The translation priming effect is generally situated at the semantic level (e.g. Grainger & Frenck-Mestre, 1998; e.g. Grainger & Frenck-Mestre, 1998, see also the General Discussion Section of Chapter 4). However, because semantically related words are often also associatively related, it is also possible that such facilitative effects are due to associations between lexical representations, without semantic mediation. This issue could be investigated by looking for a cross-language priming effect of associatively (because they make up a compound noun for example), but not semantically related words. For example, if such priming effects are of an associative nature, it should be possible to prime the English (L2) word LEAD with the Dutch prime *pot* (pot-lood) (for a monolingual equivalent, see Davies, 1998).

Finally, at a more general level and following the modeling section above, I firmly believe there should be increased attention to computational modeling of the interaction between orthographic, semantic and phonological representations in bilingual language processing, similar to proceedings in the monolingual domain (see also the evolutions in the BIA+ model, Dijkstra & Van Heuven, 2002). Up to now, the BIA model has done a very good job in dealing with orthographic representations, leading to several - intuitively implausible - important empirical results, and therefore increasing our understanding of bilingual visual word recognition. We believe classical box-arrow models that make explicit assumptions with respect to the interaction between several levels of representation, such as the RHM, should follow this example. At present, the different modules of the RHM basically are black boxes. Computational modeling should increase the model's clarity and level of detail, so that much more explicit assumptions and predictions can be made with respect to bilingual language processing at different representational levels. This should allow to integrate the results obtained in the present dissertation with the studies supporting the RHM, and to explain for example influences of word variables on translation performance.

REFERENCES

Altarriba, J., & Mathis, K. M. (1997). Conceptual and Lexical Development in Second Language Acquisition. *Journal of Memory and Language, 36*(4), 550-568.

Bloem, I., & La Heij, W. (2003). Semantic Facilitation and Semantic Interference in Word Translation: Implications for Models of Lexical Access in Language Production. *Journal of Memory and Language, 48*(3), 468-488.

Brysbaert, M., Van Dyck, G., & Van de Poel, M. (1999). Visual Word Recognition in Bilinguals: Evidence From Masked Phonological Priming. *Journal of Experimental Psychology: Human Perception and Performance, 25*(1), 137-148.

Brysbaert, M., Van Wijnendaele, I., & Duyck, W. (2002). On the temporal delay assumption and the impact of non-linguistic context effects. *Bilingualism: Language and Cognition, 5*(3), 199-201.

Davies, P. D. (1998). Automatic Semantic Priming: the Contribution of Lexical- and Semantic-Level Processes. *European Journal of Cognitive Psychology, 10*(4), 389-412.

de Groot, A. M. B. (1992). Determinants of Word Translation. *Journal of Experimental Psychology: Learning, Memory and Cognition, 18*(5), 1001-1018.

de Groot, A. M. B. (1993). Word-type Effects in Bilingual Processing Tasks. Support for a Mixed-Representational System. In Schreuder, R. and Weltens, B. (Eds.), *The Bilingual Lexicon: Vol. 6. Studies in Bilingualism.* (pp. 27-57). Amsterdam, Netherlands: John Benjamins Publishing Company.

de Groot, A. M. B., Dannenburg, L., & van Hell, J. G. (1994a). Forward and Backward Word Translation by Bilinguals. *Journal of Memory and Language, 33*(5), 600-629.

de Groot, A. M. B., Dannenburg, L., & van Hell, J. G. (1994b). Forward and Backward Word Translation by Bilinguals. *Journal of Memory and Language, 33*(5), 600-629.

Dijkstra, A., Grainger, J., & Van Heuven, W. J. B. (1999). Recognition of Cognates and Interlingual Homographs: the Neglected Role of Phonology. *Journal of Memory and Language, 41*(4), 496-518.

Dijkstra, A., & Van Heuven, W. (2002). The Architecture of the Bilingual Word Recognition System: From Identification to Decision. *Bilingualism: Language and Cognition, 5*(3), 175-197.

Duyck, W., & Brysbaert, M. (2002). What Number Translation Studies Can Teach Us About the Lexico-Semantic Organisation in Bilinguals. *Psychologica Belgica, 42*(3), 151-175.

Duyck, W., Diependaele, K., Drieghe, D., & Brysbaert, M. (2004). The Size of the Cross-Lingual Masked Phonological Priming Effect Does not Depend on Second Language Proficiency. *Experimental Psychology, 51*(2), 1-9.

Duyck, W., Szmalec, A., Kemps, E., & Vandierendonck, A. (2003). Verbal Working Memory Is Involved in Associative Word Learning Unless Visual Codes Are Available. *Journal of Memory and Language, 48*(3), 527-541.

Francis, W. S., Augustini, B. K., & Saenz, S. P. (2003). Repetition Priming in Picture Naming and Translation Depends on Shared Processes and Their Difficulty: Evidence From Spanish-English Bilinguals. *Journal of Experimental Psychology: Learning Memory and Cognition, 29*(6), 1283-1297.

Frost, R. (1998). Toward a Strong Phonological Theory of Visual Word Recognition: True Issues and False Trails. *Psychological Bulletin, 123*(1), 71-99.

Gollan, T. H., Forster, K. I., & Frost, R. (1997). Translation Priming With Different Scripts: Masked Priming With Cognates and Noncognates in Hebrew-English Bilinguals. *Journal of Experimental Psychology: Learning, Memory and Cognition, 23*(5), 1122-1139.

Grainger, J., & Frenck-Mestre, C. (1998). Masked Priming by Translation Equivalents in Proficient Bilinguals. *Language and Cognitive Processes, 13*(6), 601-623.

Jared, D., & Kroll, J. F. (2001). Do Bilinguals Activate Phonological Representations in One or Both of Their Languages When Naming Words? *Journal of Memory and Language, 44*(1), 2-31.

Kroll, J. F., & de Groot, A. M. B. (1997). Lexical And Conceptual Memory in the Bilingual: Mapping Form to Meaning in Two Languages. In de Groot, A. M. B. and Kroll, J. F. (Eds.), *Tutorials In Bilingualism: Psycholinguistic Perspectives.* (pp. 201-224). Mahwah, NJ: Lawrence Erlbaum Publishers.

Kroll, J. F., & Stewart, E. (1994). Category Interference in Translation and Picture Naming - Evidence for Asymmetric Connections Between Bilingual Memory Representations. *Journal of Memory and Language, 33*(2), 149-174.

La Heij, W., Hooglander, A., Kerling, R., & Van der Velden, E. (1996). Nonverbal Context Effects in Forward and Backward Word Translation: Evidence for Concept Mediation. *Journal of Memory and Language, 35*(5), 648-665.

Schoonbaert, S., Duyck, W., and Brysbaert, M. (2004). *Backward Translation Priming Interacts with Word Concreteness*. Manuscript in preparation.

Tan, L. H., & Perfetti, C. A. (1997). Visual Chinese Character Recognition: Does Phonological Information Mediate Access to Meaning? *Journal of Memory and Language, 37*(1), 41-57.

Tzelgov, J., Yehene, V., Kotler, L., & Alon, A. (2000). Automatic Comparisons of Artificial Digits Never Compared: Learning Linear Ordering Relations. *Journal of Experimental Psychology: Learning, Memory and Cognition, 26*(1), 103-120.

Van Hell, J. G., & de Groot, A. M. B. (1998a). Conceptual Representation in Bilingual Memory: Effects of Concreteness and Cognate Status in Word Association. *Bilingualism: Language and Cognition, 1*, 193-211.

Van Hell, J. G., & de Groot, A. M. B. (1998b). Disentangling Context Availability and Concreteness in Lexical Decision and Word Translation. *Quarterly Journal of Experimental Psychology, 51A*, 41-63.

van Hell, J. G., & de Groot, A. M. B. (1998c). Disentangling Context Availability and Concreteness in Lexical Decision and Word Translation. *Journal of Experimental Psychology, 51A*, 41-63.

Van Wijnendaele, I., & Brysbaert, M. (2002). Visual Word Recognition in Bilinguals: Phonological Priming from the Second to the First Language. *Journal of Experimental Psychology: Human Perception and Performance, 3*, 619-627.

von Pein M., & Altarriba J. (2003). *The Development of Linguistic Knowledge in Naïve Learners of American Sign Language*. Manuscript submitted for publication.

Zhou, X. L., & Marslen-Wilson, W. (1999). Phonology, Orthography, and Semantic Activation in Reading Chinese. *Journal of Memory and Language, 41*(4), 579-606.